

# **LOGIC AND INTELLIGIBILITY**

by

James Jonathan Pearson

B.A., University of Oxford, 2003

Submitted to the Graduate Faculty of the  
Kenneth P. Dietrich School of Arts and Sciences in  
partial fulfillment of the requirements for the degree of  
Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

Dietrich School of Arts and Sciences

This dissertation was presented

by

James Jonathan Pearson

It was defended on

August 24, 2011

and approved by

Anil Gupta, Distinguished Professor, Philosophy Department

Stephen Engstrom, Professor, Philosophy Department

Steven Awodey, Professor, Philosophy Department, Carnegie Mellon University

Dissertation Director: Thomas Ricketts, Professor, Philosophy Department

Copyright © by James Jonathan Pearson

2011

## LOGIC AND INTELLIGIBILITY

James Pearson, Ph.D.

University of Pittsburgh, 2011

All inquirers must have a grasp of implication and contradiction which they employ to structure their investigations. Logical ability is thus some kind of prerequisite for cognition. My dissertation scrutinizes this relationship and argues that different ways of understanding it underlie a deep debate about naturalism and the objectivity of our knowledge.

Frege's dismissal of logical aliens as mad exposes his conviction that logical ability is exhibited in our practice of demonstrative reasoning, and is a *constitutive* necessary condition for cognition. By denying the existence of an independent standard for objective truth that a codification of inferential principles must meet, Frege avoids logical "sociologism" (under which the validity of inferential principles is identified with their agreement with our practice).

Quine objects to ascribing a "pre-logical mentality" in radical translation, but only because doing so would represent one's interlocutor as affirming something one finds obviously false. Under his *naturalism*, the logician is guided by usefulness to ongoing empirical inquiry, not the search for the constitutive prerequisites of thinking. I argue that the properly-understood naturalist excises various skeptical attacks from epistemology.

Davidson recovers a privileged status for logic as central in the theories of truth that are necessary to interpret another as—and also to be—a cognizer. Under his *humanism*, it is only

through interpreting others that one can grasp the objective/subjective contrast and acquire beliefs that are properly about the world. We do not exhibit our grasp of objective truth by engaging in a practice informed by logic, but by interpreting others who are engaged with, and through, us in such a practice.

Despite initial appearances, naturalism and humanism are not incompatible positions. After examining Quine's "sectarian" and Davidson's "ecumenical" attitude to the truth of empirically equivalent theories, I endorse ecumenism about their metaphilosophical disagreement. By renouncing a proprietary attitude to truth, this particular form of tolerance avoids the fragmentation of philosophy into distinct, yet totalizing, and hence warring, programs.

## TABLE OF CONTENTS

<b>Preface.....</b>	<b>viii</b>
<b>I. INTRODUCTION.....</b>	<b>1</b>
<b>II. CHAPTER 1: Frege’s Constitutive View and Our Certain Knowledge of Logic.....</b>	<b>6</b>
1. Burge’s Frege: A Pragmatic Euclidean.....	7
2. Burge’s Evidence for attributing a “Pragmatic” Justification to Frege.....	16
3. Weiner’s Frege: A Logocentric Elucidator.....	26
4. Madness and Logical Aliens.....	40
5. Sullivan’s Rejoinder.....	62
Bibliography.....	71
<b>III. CHAPTER 2: Quine’s Naturalist Conception of Logic: Excising Skepticism from Epistemology.....</b>	<b>75</b>
1. Naturalism, Logic, and Cognition: The Denial of a Pre-Logical Mentality.....	76
2. Barry Stroud and Skepticism About the External World.....	87
3. Rudolf Carnap and Skepticism About Successful Communication.....	98
4. Gary Ebbs and Skepticism About The Use of Our Words.....	116
Bibliography.....	132
<b>IV. CHAPTER 3: Davidson’s Humanism: Truth Only Through Interpretation.....</b>	<b>137</b>
1. Radical Interpretation and Reading Logic Into Language.....	138
2. The Charge of Verificationism and Davidson’s Transcendental Argument.....	156
3. The Humanist Approach.....	168
Bibliography.....	176
<b>V. CHAPTER 4: Disagreement in Philosophy: The Case of Naturalism Versus Humanism.....</b>	<b>180</b>

1. Davidson's Skeptical Attacks Against Quine.....	181
2. Truth, Objectivity, and Solipsism.....	189
3. Resolving Philosophical Disagreement.....	203
Bibliography.....	215
<b>VI. AFTERWORD.....</b>	<b>218</b>

## **Preface**

I would like to begin by thanking the members of my dissertation committee: Thomas Ricketts, Anil Gupta, Stephen Engstrom, and Steven Awodey. Their interest in the nascent stages of my project has helped to keep me motivated during the years that have followed. I owe special thanks to my dissertation advisor, Tom. In our many conversations, Tom deftly balanced his textual mastery and philosophical acumen with the pedagogical goal of helping me to find my own voice. His trenchant criticisms of my writing have taught me how much doing good philosophy, and good history of philosophy, is a matter of carefully finding the words to express one's thoughts. I hope one day that I can be to my students what he has been to me. Anil, as the second reader of my dissertation, has also been exemplary, promptly providing extremely helpful feedback on any chapter drafts I sent him.

I also wish to thank the philosophy faculty at the University of Pittsburgh who fueled my interest in various areas during my years of coursework. Three directed studies in particular—one with Bob Brandom on philosophy of language, another with Anil on philosophy of logic, and finally one with Tom on Carnap and Quine—proved instrumental when choosing my dissertation topic. Bob taught me the importance of writing something every day; Anil taught me the value of encapsulating the arguments in what I had read; and Tom taught me how exciting history of philosophy can be.

Perhaps it was inevitable that in thinking for so long about logical aliens there would be times that it seemed I had become one. For their help in getting me through these moments, I would like to thank three people in particular. The innumerable discussions that I have had with Kyle Stevens, Evgenia Mylonaki, and Hille Paakkunainen have yielded philosophical insights and encouragement at different stages of the project. Beyond helping me to acknowledge the



importance of triangulation, Kyle has painstakingly edited successive drafts and greatly improved my writing. Conversations with Evgenia were always passionate and engrossing, and reminded me to enjoy my work. Hille's extensive, careful feedback, particularly on chapter one, helped me to keep in mind the connections between history and contemporary philosophy. I have tremendous respect for these three thinkers, and owe each a debt of gratitude. I would also like to thank the many Pittsburgh graduate students with whom I have discussed different parts of this dissertation: Dan Addison, Brett Caloia, Kim Frost, Sasha Newton, Tyke Nunez, Jamsheed Siyar, Shawn Standefer, Markos Valaris, Tim Willenken, and Christine Young. I also thank Michael Hoffman for his assistance in designing the figures in chapter one.

I dedicate this dissertation to the memory of my sister, Rebecca Pearson.

## **Introduction**

To structure investigations, each inquirer must have some grasp of implication and contradiction. One who exhibits no such grasp is not recognizable as an inquirer. Thus, logic is a kind of prerequisite for cognition. Other minds become intelligible once we perceive them reasoning according to logical standards; logic makes thoughts, reasons, and arguments intelligible *as* thoughts, reasons, and arguments.

In this dissertation, I argue that different ways of understanding how logic makes cognition intelligible underlies a deep debate about the objectivity of knowledge. In twentieth-century analytic philosophy, this epistemological issue has been bound to the methodological problem of how philosophy ought to be done. By scrutinizing how different conceptions of the relationship between logic and cognition subtend different philosophical approaches, I will illuminate the recent history of analytic naturalism and its malcontents.<sup>1</sup> This allows me to contribute to recent interest in the epistemology of disagreement by considering the case of intractable disagreement in philosophy.

In chapter one, I argue that one understanding of logic's relationship to cognition informs Gottlob Frege's conception of logic. In defending his view of the autonomy and objectivity of logic against psychologism, Frege imagines logical aliens (beings who sincerely deny a logical law) and dismisses them as "mad." This dismissal exposes Frege's conviction that logical ability is exhibited in our practice of demonstrative reasoning. Frege thinks that this ability is a *constitutive* necessary condition for cognition—indeed, for the capacity for thought. I argue that Frege conceives of his Begriffsschrift as an explicit codification of the inferential principles

---

<sup>1</sup> Throughout this dissertation, I use "cognition" to refer to a cluster of epistemic capacities, including our capacities to judge, to infer, to justify, and to think. Each of the philosophers I discuss elaborate our cognitive abilities in different ways. Looking at their different accounts of the relationship between logic and cognition is, and at the same time, to look at their different accounts of logic and their different accounts of cognition.

implicit in our demonstrative practice, and so holds that any inquirer who understands his formulations of these principles should embrace them as certain. Tyler Burge is therefore incorrect about the order of justification for Frege's notation. Burge thinks that the principles of Begriffsschrift are pragmatically justified because instances of our demonstrative reasoning can be formally reconstructed using the notation. Yet, any such pragmatic certification would necessarily fall short of the certainty that Frege demands.

On my reading, the Fregean logician formulates the principles of inference implicit in our demonstrative practice. Peter Sullivan argues that such a view amounts to what I call a "sociologism" that (analogously to the psychologism Frege rejects) is incompatible with the objectivity of logic and truth. This view, Sullivan claims, forces one to identify the validity of inferential principles with their agreement with our practice. But, Sullivan's objection presupposes the existence of an independent standard for objective truth that a purported codification of inferential principles must meet. Against him, I argue that there is no such standard. For Frege, the principles of inference are not valid because they are implicit in our practice, but are implicit in our practice because they are valid. Frege understands recognition of the laws of logic (which he also calls the laws of truth) to be constitutive of an inquirer's grasp of objective truth.

Frege's view of logical ability as a constitutive necessary condition for cognition provides a helpful backdrop for contrasting W.V. Quine's and Donald Davidson's conceptions of logic and objectivity. Frege dismisses logical aliens as mad. Similarly, Quine maintains that translating another person as affirming contradictions is evidence of a mistake in our translation manual. For Frege, the aliens fail to meet the constitutive standard for the attribution of cognition: recognition of the logical laws as immutable, eternal "boundary stones" for thought.

Yet Quine's objection to ascribing a "pre-logical mentality" is not that in so doing we represent a putative thinker affirming things that no genuine thinker could intelligibly affirm; it is simply that we represent a thinker as affirming something we find obviously false. Just as translating people consistently calling red objects "blue" is not credible, it is not credible to translate people as affirming simple contradictions. No principled difference in status distinguishes logical from physical truths. Rather, beliefs toward the periphery of our "web of beliefs" are easily given up, while obvious ones near the center (both logical and physical) are more difficult to reject. Denying logic a privileged, fixed relation to cognition means that Quine cannot adopt Frege's account of the link between our conception of objectivity and our logical ability. So how does Quine establish objectivity?

To answer this question, I unpack Quine's appeal to naturalism in chapter two. Quine holds philosophical and scientific inquiry to be one continuous, fallible enterprise. This view has often been misunderstood. Against Barry Stroud and Gary Ebbs, I show how the naturalist excises a variety of skeptical arguments from traditional epistemology. Quine's denial that the philosopher—including the logician—can occupy a supra-scientific position stems from his rejection of "cosmic exile," the notion that it is intelligible to look down from above and evaluate all of our beliefs at once. Instead, Quine claims that we can only question our beliefs from the side, on the basis of other beliefs that are not then and there at issue. He esteems scientific methodology as the best method inquirers have for such sideways evaluation. It has proven to result in theories of reality with unparalleled predictive and explanatory power. Naturalist inquirers show themselves accountable to the objective standard of truth by recognizing their own fallibility, continually refining (and accepting others' refinements of) their theories. The codification of logic developed by a naturalist logician is an attempt to explicate "implication,"

the relation between sentences to which scientists informally appeal when reasoning. Such a codification allows ontological philosophers to regiment our theories and delineate the categories to which we are currently committed. The principle guiding the naturalist logician is usefulness to ongoing empirical inquiry, not the search for constitutive prerequisites of thinking.

Davidson refuses Quine's naturalism. In chapter three, I examine his argument that interpreters form an "epistemological triangle" with the world. To emphasize the importance intersubjectivity has to this account, I designate Davidson's view "humanism." Davidson's humanism recovers a privileged status for logic as central in theories of truth that are necessary to interpret another as—and also to be—a cognizer. I argue, against Simon Blackburn, that Davidson's argument against uninterpretable languages is not tacitly verificationist, but transcendental. Not only does Davidson think that developing a truth theory allows us to interpret another person's utterances as meaningful; he thinks that it is only through engaging in such interpretive acts, grounded in quantificational logic, that one can become aware of the objective/subjective contrast and so come to have beliefs that are about the world. The social condition humanism imposes upon cognition distinguishes it from constitutivism. We do not exhibit our grasp of objective truth by engaging in a practice informed by logic, but by interpreting others who are engaged with us in such a practice.

In chapter four, I evaluate how we should view the dispute between Quine and Davidson. By separating unconvincing ancillary considerations from the core of Davidson's objections to Quine, I expose that attributing scheme/content dualism to Quine, as Davidson does, is to argue that he lacks an adequate conception of objectivity. This charge stands despite the various modifications that Quine makes to his views under pressure from Davidson's sustained critique. Whereas Quine can turn to naturalism to convincingly respond to skeptical attacks that others

have mounted against him, Davidson's more nuanced attack reveals that in refusing to follow the humanist in externalizing epistemology, the naturalist cannot rule solipsistic threats unintelligible.

Despite initial appearances, naturalism and humanism are not incompatible positions at most one of which is true. I close by tentatively endorsing a version of metaphilosophical tolerance which, by forcing us to occupy that perspective which our current inquiry demands, avoids the fragmentation of philosophy into distinct totalizing programs. In my view, the choice between adopting naturalism and humanism depends upon the details of the philosophical problem currently occupying us, a lesson that can be generalized to other cases of intractable philosophical disagreement.

## **Frege's Constitutive View and Our Certain Knowledge of Logic**

According to Tyler Burge, Gottlob Frege believes that the wide applicability of his Begriffsschrift in mathematics and science justifies holding it true and holding it to truly codify the objective laws of logic.<sup>1</sup> This “pragmatic” reading of Frege’s epistemology of logic is mistaken. A consequence of Frege’s constitutive account of the relationship between logic and cognition is that logical laws neither need nor admit of justification. For Frege, we exhibit certain knowledge of logic by engaging in demonstrative reasoning.

After presenting Burge’s interpretation in section one, I argue in section two that he fails to provide it with adequate textual support. In section three, I explain why the passages that Burge cites better comport with Joan Weiner’s deflationary interpretation of Frege’s conception of logic and epistemology. By articulating the role that Frege’s constitutive view plays in his account of the distinctive unintelligibility of a particular species of logical aliens in section four, I argue that, unlike Burge’s, a deflationary interpretation can capture the sense in which Frege understands engaging in our inferential practice to be a requirement of cognition. In section five, I consider Peter Sullivan’s rejoinder that so interpreting Frege implausibly commits him to what I shall call “sociologism” about logic. Although Sullivan’s objection is not convincing, it highlights Frege’s peculiarly static conception of our inferential practice and reveals that his account of logic’s role in communication is underdeveloped.

---

<sup>1</sup> Throughout this chapter, I shall use “*Begriffsschrift*” to refer to Frege’s 1879 book and “Begriffsschrift” to refer to his logical notation.

## **1. Burge's Frege: A Pragmatic Euclidean**

From his recent work, we can extract four theses to which Burge thinks Frege is committed:

(B1) *Objectivity as Ontological Independence*: Thoughts—including the laws of logic—are “objective” in the sense that they *exist* wholly independently of human activity.<sup>2</sup>

(B2) *Platonism about Thoughts is an Explanatory Theory*: The existence of thoughts in a “third realm” explains the possibility of intersubjective communication.<sup>3</sup>

(B3) *The Begriffsschrift admits of semantic “bottom-up” justification*: Arguments using a truth predicate for purposes of semantic ascent can justify that the syntactically-specified inference rules of the formalism are truth-preserving.<sup>4</sup>

(B4) *The Begriffsschrift admits of pragmatic “top-down” justification*: The application of the formalism to informal scientific and mathematical reasoning can justify taking its axioms to be true, taking its inference rules to be truth-preserving, and taking it to codify the logical principles implicit in colloquially stated demonstrative argumentation.<sup>5</sup>

I do not think that Frege is committed to any of these theses, and attributing them badly distorts his conception of logic and epistemology. Although I shall not explicitly criticize Burge's attribution of B1-B3 to Frege in this chapter, in this section I want to discuss how B1-B4 figure in his overall interpretation, before arguing that Burge's textual evidence for attributing B4 to Frege is not convincing in the next section.<sup>6</sup> There are three reasons for this exegetical work. Firstly, B4 is motivated by the epistemological problems that Burge interprets Frege's project as facing. Secondly, exploring Burge's position in detail will allow me to contrast it with Weiner's, thereby showing why scholars like Dirk Greimann are wrong to count their interpretations as

---

<sup>2</sup> Burge (2005, 145; 2005, 305).

<sup>3</sup> Burge (2005, 304-305).

<sup>4</sup> Burge (2005, 333-337).

<sup>5</sup> Burge (2005, 342; 2005, 355).

<sup>6</sup> For criticisms of attributing B1 and B2 to Frege, see Thomas Ricketts (1986, 1996) and Weiner (1995a, 1995b). For criticisms of attributing B3 to Frege, see Ricketts (1996) and Weiner (2005, 2008).



terminological variants (2008, 406). Finally, this work positions us to see how Burge's reading of Frege's constitutive view is incompatible with Frege's texts.

Frege writes that, although the process of thinking is subjective, the thoughts that individual inquirers think are "objective," "independent of those who judge," and are true or false regardless of whether any person judges them to be so (T, 73; FA, §26; GGZ1, xvi). Were thoughts subjective, Frege claims, two scientists could not contradict each other because "each...[individual would only be] concerned with [the] contents of his own consciousness" and so "it would really be idle to dispute about truth" (T, 69). Emphasizing Frege's apparently Platonist claim that thoughts "belong" to a "third realm" (T, 69), Burge interprets his attribution of "objectivity" to thoughts ontologically (B1). Since Frege views his argument against the subjectivity of thoughts to be a *reductio*,<sup>7</sup> Burge takes Platonism to be a substantive theory meant to explain intersubjective communication (B2).<sup>8</sup> Burge's Frege thinks of scientific inquiry as the quest to discover which Platonic thoughts are true.

Scientists inquire by inferring truths on the basis of others. The science of logic studies the relation of inference upon which such inquiry depends. Burge views Frege's approach to logic as belonging to the Euclidean tradition. Euclid used an axiomatization to make the special science of geometry rigorous. Euclid's axioms express basic truths about simple geometric vocabulary that can be used to define and prove theorems about more complex geometrical figures. Since one must appeal to principles of inference to justify inferring theorems on the basis of axioms, Euclid's method relies upon logic. This complicates Burge's Frege's project of

---

<sup>7</sup> See, for example, (CP, 368), where Frege writes that he is certain that he *does* communicate with other judges.

<sup>8</sup> Burge resists affirming that Frege's Platonist ontological theory is properly *prior* to his logical theory, instead claiming that, for Frege, logic and ontology are "mutually entangled" (1992, 644n). But this is enough to situate his reading in opposition to Weiner's, who, following Ricketts (1986), argues that Frege's conception of ontological categories *supervene* on his conception of logical categories (1995b, 376-378). It is the supervenience claim that prohibits interpreting Frege as committed to a substantive ontological theory whose expression violates logical categories, as I shall show in section three.

developing an axiomatization *of* logic, for, in addition to identifying some Platonic thoughts and expressing them as logical axioms, it must make explicit the logical rules of inference that license the deduction of further truths from those axioms and given non-logical truths.

Let us begin with the logical axioms. How must Burge's Frege view them? Frege writes that the Euclidean axioms govern our thoughts about "spatially intuitable" domains (FA, §14). In contrast, axioms of logic ought to govern our thoughts in every domain, and so, must abstract away from the content of special sciences. Frege concludes that logical axioms must express *maximally* general truths, capable of being articulated using only topic-universal notions necessary for the expression of knowledge in any special science.<sup>9</sup> Similarly, logical rules of inference must express valid transitions between statements in any science, not statements of a particular science. However, in the context of the Platonism Burge attributes to him, Frege's project raises what has come to be known as the demarcation problem. If maximal generality is going to serve as a *criterion* by which logical truths are distinguished from non-logical truths in the third realm, we need an account of how to identify topic-universal vocabulary. For otherwise, why think that Frege's Begriffsschrift, as opposed to some stronger (or weaker) axiomatization with a richer (or poorer) vocabulary counted "topic-universal," accurately demarcates the objective principles of demonstrative inference?

The demarcation problem questions why we should think a given formalism codifies logic. But there is a second epistemological problem for Burge's Frege: why should we think the axioms of a formalism *true*, and its rules *truth-preserving*? Although Frege tells us that we can justify a statement by proving it from primitive truths (FA, §3), logical axioms are foundations and do not admit of proof. How, then, are axioms justified?

---

<sup>9</sup> Frege writes, "logic is the science of the most general laws of truth" (PW, 128). For discussion, see Ricketts (1986b).

Frege, in keeping with the Euclidean tradition, calls the basic laws of his Begriffsschrift—both its axioms and its rules of inference—*selbstverständlich* or *einleuchtend* or *unmittelbar klar*, all of which are usually translated as “self-evident” (FA, §5; §90; PW, 39; 208-210). He never defines these terms. They seem to indicate that Frege thinks logical laws obvious. Once one understands an axiom, one is wholly confident that it is true. But “obviousness” and “confidence” are matters of psychology. Frege cannot mean to invoke *obviousness* as a justification for our knowledge of primitive truths as true, for then—like the psychological logicians against whom he tirelessly crusades—he would be allowing psychology to cloud his logical investigation.<sup>10</sup> One may doubt what another judge finds obvious. One’s confidence about a putative truth may be eroded. Nor does attributing “self-evidence” in the sense of obviousness to his codification answer the demarcation problem. If he were to conclude that any general law obvious to him was logical, then his demarcation would rely upon his own psychology.<sup>11</sup> What role is left for self-evidence in Frege’s epistemology of logic?<sup>12</sup>

Burge credits Frege with a nuanced understanding of self-evidence (2005, 346-354). He claims Frege believed that the psychological and subjective impression of obviousness was distinct from the non-psychological and objective property of self-evidence, in the sense of being *self-justifying*, enjoyed by the true laws of logic. Just as theoretical physicists attempt to formulate the laws which govern motion, Burge thinks that the Fregean logician attempts to formulate the laws of inference. Yet neither inquirer is, nor thinks herself, infallible. The

---

<sup>10</sup> Frege explicitly rejects the possibility that psychological facts can justify logical truths (GGZ1, xvii). As Ludwig Wittgenstein would later put it, “if the truth of a proposition does not *follow* from the fact that it is self-evident to us, then its self-evidence in no way justifies our belief in its truth.” (2002, 5.1363).

<sup>11</sup> Wittgenstein observes: “it is remarkable that a thinker as rigorous as Frege appealed to the degree of self-evidence as the criterion of a logical proposition” (2002, 6.1271). I do not think this objection is decisive against Frege, because the demarcation problem is irrelevant to his actual project. I discuss this point in the following section.

<sup>12</sup> Frege’s claim that “we cannot accept a thought as an axiom if we are in doubt about its truth” (PW, 205) might suggest that he thought *rationally indubitable* obvious truths self-evident. But this modification does not provide a non-psychological answer to either epistemological problem, since the truths which appear rationally indubitable *to Frege* are not necessarily truths that *are* rationally indubitable.

Fregean logician's confidence in her codification is partly justified by her subjective intuitions about its obviousness. But her intuitions about its obviousness do not establish its truth or logicity. She recognizes that what she holds to be self-evidently true might turn out to be false, and that what she holds to be a primitive logical truth might turn out to be true but in need of proof on the basis of still more primitive (and actually logical) truths.

Robin Jeshion's recent development of Burge's interpretation clarifies how this nuanced conception of self-evidence shapes the epistemological problems of justifying one's codification as true, and as logical (2001, 2004). Recall that Frege uses a variety of German terms (*selbstverständlich*, *einleuchtend*, *unmittelbar klar*) all of which are translated as "self-evident." Although Jeshion admits that Frege never gives these terms different technical meanings, and that no putative explication fits all of his texts precisely,<sup>13</sup> she offers the following explication of Frege's view:

A true proposition *p* is *selbstverständlich* if and only if *p* is foundationally secure, yet *p*'s truth is not grounded on any other truth.<sup>14</sup>

A true proposition *p* is *self-evident* [*einleuchtend*, *unmittelbar klar*] if and only if clearly grasping *p* is sufficient and compelling basis for recognition of *p*'s truth.<sup>15</sup>

Jeshion takes neither definition to be subjective or psychological.<sup>16</sup> She characterizes Frege's logicist project as the "search for foundational arithmetical axioms...[that are] both *selbstverständlich* and self-evident," and believes that, since all self-evident propositions are *selbstverständlich*, Frege "relied on judging propositions to be self-evident as part of his method

---

<sup>13</sup> Jeshion (2001, 953). Burge also confesses that he has not "found consistent differences in Frege's usage" of these terms (2005, 346).

<sup>14</sup> Jeshion (2001, 949). Jeshion writes that "A *ground* of a true proposition indicates from whence the proposition's truth derives" (2001, 945). Thus, the truth of a *selbstverständlich* thought does not derive from any other thought.

<sup>15</sup> Jeshion (2001, 953). Jeshion leaves *selbstverständlich* in German because she can find no clear correlate in English, but accepts the standard translation of *einleuchtend* as "self-evident." She believes that Frege's use of *unmittelbar klar* (immediately clear) is also intended to convey "self-evidence" in her technical sense.

<sup>16</sup> This is clear for Jeshion's explication of "*selbstverständlich*," but less clear for her explication of "self-evident," since its mention of "grasping" implies a grasper. But in Jeshion's view, the implied grasper is an *idealized, fully rational* grasper, not a particular human grasper.

for identifying a foundation for arithmetic” (2001, 939).<sup>17, 18</sup> But it is necessary, in turn, to rely on subjective obviousness as a guide to judging propositions self-evident (967). Obviousness is a fallible guide, because not all propositions one finds obvious are self-evident, and one may not find all self-evident propositions obvious (if one lacks the requisite understanding of component concepts). Nevertheless, Jeshion thinks obviousness is “normally necessary” for identifying primitive truths.<sup>19</sup>

Under Jeshion’s explication, Frege’s invocation of “self-evidence” is not psychological but normative. It concerns the immediate recognition of a thought as true by an *ideally* rational person who possesses an *entirely clear* understanding of its component concepts. Logic is demarcated by what is *actually* self-evident to, as Burge puts it, an “ideal rational mind” (2005, 350), not what *seems* obvious to particular minds, such as Frege’s own. Unlike the author of the *Tractatus*, whom Jeshion characterizes as “preoccupied” with certainty (2001, 972), Frege “never took particular proposed foundations as certain” (973)—even his own proposed codification of logic, the *Begriffsschrift*.<sup>20</sup> However, Frege was extremely confident that he had successfully codified logic (GGZ1, xxvi). If Burge and Jeshion are correct that Frege thought his method (insofar as it relied upon the perception of obviousness) fallible, then he owes any reader

---

<sup>17</sup> Frege, according to Jeshion, believes that the converse implication does not hold. An example is the equation Frege gives in (FA, §5):  $135664 + 37863 = 173527$ . All such true arithmetical equations are really disguised definitions, and so, are *selbstverständlich*. They lack the generality required to stand as axioms of arithmetic, but are nevertheless foundationally secure in and of themselves (2001, 950). Yet if, like this equation, they are sufficiently complex, they are not self-evident and thus require proof. To Jeshion’s Frege, “if one grasps a self-evident proposition *p* yet fails to thereby recognize its truth, the failure to do so implies a deficiency in conceptual understanding or rationality” (956). Since a rational agent can understand such an equation and yet (prior to checking it for herself) refrain from endorsing or rejecting it, the equation is not self-evident.

<sup>18</sup> Although Jeshion does not develop this idea explicitly in her paper, it seems that if she is correct, Frege would view axioms in the ideal axiomatization of a special science to be *selbstverständlich* but not self-evident (PW, 205). An example might be Henry’s Law in the science of gases. Henry’s law is itself foundationally secure, a general expression of the natural proportion of pressure and concentration of gases at constant temperatures; but recognizing its truth requires a good deal of empirical investigation.

<sup>19</sup> It is only “normally” necessary because we are also guided by the pragmatic considerations that I will discuss in a moment. For some inquirers, these pragmatic considerations could be methodologically sufficient by themselves.

<sup>20</sup> Jeshion complains that Wittgenstein’s considerable influence has caused Frege’s nuanced Euclideanism to have been missed (2001, 973).

who does *not* find the Begriffsschrift obvious a justification for his confidence in its truth and logicality.

Burge claims Frege offers two sorts of considerations that are meant to transmit his confidence to his reader. The first are the “semantic claims” he makes when introducing each individual rule and axiom of his Begriffsschrift (2005, 328). Burge correctly notes that, since Frege thought his system codified the objective laws of logic, he could not have thought that its axioms admitted of proof on the basis of more fundamental considerations. So the purpose of the semantic claims cannot be to support (or justify) the *truth* of the axioms (330). Instead, they serve to bring out the self-evidence of the axioms to a readership unfamiliar with Frege’s novel notation. In contrast, in §14 of *Grundgesetze*, Burge interprets Frege as developing a genuine semantic *argument* to justify the soundness of the Begriffsschrift inferential rule of *modus ponens* (2005, 332). By reflecting on the *meanings* of the material conditional, the horizontal, and the judgment stroke, one ought to recognize that the transition licensed by the notation’s inference rule from true sentences whose syntactic forms are a conditional and its antecedent to the sentence whose syntactic form is the conditional’s consequent *is* a truth-preserving inference (B3).<sup>21</sup>

Yet while such bottom-up semantic considerations might plausibly transmit confidence in the *truth* of individual Begriffsschrift laws, they do little to transmit confidence that the Begriffsschrift codifies the principles of demonstrative inference. We now learn predicate logic *as* logic in introductory classes, but Frege’s readers would find many of his innovations (such as using the artificial material conditional rather than the syllogistic forms to express universal categoricals) obscure. In fact, readers put off by the Begriffsschrift’s lack of familiarity might

---

<sup>21</sup> Burge acknowledges that this semantic bottom-up justification of the soundness of individual inferential rules is unsystematic by modern lights (331). He finds the second set of considerations Frege offers to be his most significant contribution to the Euclidean tradition.

not find Frege's semantic considerations compelling due to lack of comprehension or diligence.

To meet this problem, Burge reads Frege as offering "pragmatic" considerations as a top-down

"justification" for the Begriffsschrift (B4):

What is original about his position is not his view that a thought might be self-evident but not seem self-evident—self-evident but not obvious to an individual. It is not his idea that subjective obviousness or subjective unobviousness might submit to reversal through deeper conceptual development and understanding. *What is original is his integration of these traditional views with his deep conception of what goes into adequate understanding.* This conception rests on his method of finding logical structure through studying patterns of inference...[U]nderstanding logical structure derives from seeing what structures are most fruitful in accounting for the patterns of inference that we reflectively engage in...Frege's pragmatism...[plays a] role in his epistemology. [It] not only play[s] a role in accounting for understanding. [It provides] a secondary, fallible, non-demonstrative *justification*. (2005, 354-355, my emphasis)

What sort of "justification" is Burge thinking of? Pragmatic considerations cannot justify that the axioms of the Begriffsschrift are true, because (to repeat) those axioms are meant to express objective logical laws that are *selbstverständlich*. Rather, Burge writes, these considerations constitute a justification "for us" (2005, 315): a justification not only for holding the formalism to truly codify logic, but *for us to hold* the Begriffsschrift laws to be true and its rules truth-preserving.

This pragmatic justification is puzzling. Burge variously calls it "inductive" (343, 355), "probable" (343), and "fallible" (355). One might use these terms in suggesting that Frege advances an *enumerative* inductive justification for Begriffsschrift. On this view, since the proof of every theorem that we have so far considered can be plausibly captured within Frege's notation (perhaps having made explicit any enthymematic premises upon which we had been tacitly relying), we are justified in holding its axioms to be true and its rules truth-preserving. But, given Frege's criticism that formalists *erroneously* think themselves justified in holding a system to be true on the basis that it has yet to engender a contradiction (see especially FA, §96),

he would surely not have accepted this contingent judgment about the Begriffsschrift's expressive power as constituting anything worth calling a "justification."<sup>22</sup>

I think that Burge has something more sophisticated in mind. Since mathematics and the more rigorous sciences are saturated with the quantificational inferences that the Begriffsschrift makes explicit, Frege's notation plausibly captures the patterns of inference upon which working scientists depend. Reflecting upon the Begriffsschrift's palpable expressive power to formalize reasoning that the doubtful reader informally recognizes as valid ought to increase her confidence that it codifies the principles of demonstrative inference, and also that its laws are true. But our scientific theories—and the patterns of inference that we take to be legitimate—evolve with our discoveries. Burge calls pragmatic justification "fallible," then, because in the course of ongoing scientific inquiry we might discover thoughts whose contents or interdependent structure cause us to adopt (or reject) a pattern of inference which Frege's Begriffsschrift had omitted (or enshrined).<sup>23</sup>

In Burge's view, recognizing this "secondary" justification for the Begriffsschrift will pave the way for inquirers to recognize its *primary* justification: its objective self-evidence. Inquirers who *use* the notation will come to grasp its novel concepts. In time, they will find the axioms involving those concepts subjectively obvious. This subjective obviousness will give them reason to think that the axioms do indeed express thoughts that have the objective property of self-evidence. Burge thinks that pragmatic considerations also speak to the demarcation

---

<sup>22</sup> Burge is certainly aware of Frege's criticisms of formalism (2005, 332-333), so I do not think he means to attribute *this* sort of inductive justification to Frege.

<sup>23</sup> It is worth noting that the claim that a Begriffsschrift law could be rejected with scientific progress does not contradict Frege's famous insistence that the laws of logic "are boundary stones set in an eternal foundation, which our thought can overflow but never displace" (GGZ1, xvi). On the pragmatic Euclidean interpretation, the objective laws of logic *are* eternal and immutable; but Frege's Begriffsschrift is only a fallible, human attempt to codify them. It might turn out to be *incomplete* if the logician lacks a requisite concept that will only come to light in the future. And its laws might be *false* if she has erred (and here, Burge makes much of Frege's erroneous Basic Law V which led to Russell's contradiction [2005, 352-353]).



problem. Beyond linking Frege's formalism to reasoning already informally recognized as valid, inquirers' eventual judgment that the Begriffsschrift's laws are self-evident just *is* a judgment that the system expresses the objective logical laws; or, as Jeshion puts it, that the Begriffsschrift "mirrors the natural ordering of truths."<sup>24</sup>

In this section, I have explained how B1-B4 fit into Burge's pragmatic Euclidean reading of Frege. Burge's rational reconstruction of Frege's conception of logic and epistemology is both sympathetic and comprehensive. It is also mistaken. In the following section, I will argue that the textual evidence Burge claims in support of his reading is unconvincing.

## **2. Burge's Evidence for attributing a "Pragmatic" Justification to Frege**

Although Frege emphasizes the power of his Begriffsschrift to formalize explicit, gap-free proofs of informally accepted demonstrative reasoning, he gives no indication that such pragmatic considerations give inquirers grounds for thinking his axioms true, or his rules of inference truth-preserving. The rhetoric of "justification" and "holding true" does not appear in the passages Burge cites. Instead, Frege wishes to exhibit advantages of his formalism over alternatives so as to recommend its adoption by scientists. The various proofs he provides of accepted theorems within his notation are intended to informally convey its universal applicability, not to ground its truth (PW, 38). In other words, Frege is attempting to demonstrate that his notation is complete (that any informally accepted theorem can be formally

---

<sup>24</sup> Jeshion (2001, 961). Jeshion acknowledges that her metaphor of the Begriffsschrift "mirroring" the objective logical laws is imperfect, since Frege did not think that logic had a *unique* axiomatization (BGS, 29; PW 206). He thought that a logical truth could find legitimate expression as an axiom in one codification and a theorem in another, meaning, to pursue Jeshion's metaphor, that two genuine codifications of logic may not be "reflections" of each other. But I think that this imperfection should make us seriously question whether Jeshion's metaphor is appropriate at all. It tacitly supports the Platonic idea that logical truths—which stand as the ground of "the" natural ordering of truths—exist independently, and that our codification is (currently) our best attempt to "reflect" them. But if the objective logical laws cannot be uniquely mirrored, why think that they have independent existence at all? Less metaphorically, why think that they constitute an independent standard against which our codification can (and must) be evaluated?

derived within the Begriffsschrift), not sound (that any theorem formally derived within the Begriffsschrift is true). Since there is no evidence that Frege intended pragmatic considerations to play a justificatory role, these passages cast doubt upon the epistemology of logic Burge attributes to him. In the next section, I shall turn to Weiner's more persuasive account of how Frege conceived of his project.

Burge bases his reading on three texts: Frege's unpublished article comparing his Begriffsschrift to Boole's logical calculus, late sections of *Foundations of Arithmetic*, and the introductory remarks to the first volume of *Grundgesetze*. He begins by noting that Frege repeatedly adverts advantages of his Begriffsschrift when introducing it, rather than stating that its basic principles are self-evident (2005, 338; 342-343). Frege writes that his novel use of courses-of-values of functions is a "vital advance" not just because he uses it to define Number, but because it allows a "far greater flexibility" in the expression of derivations (GGZ1, xi-x). He claims that his introduction of truth values makes "[everything] simpler and sharper" (x). Compared to his Begriffsschrift, he complains that the "Boolean formula-language only represents a part of our thinking" (PW, 35). Frege thus appeals to considerations of *flexibility*, *simplicity*, and *generality* as advantages of his Begriffsschrift. However, Burge mistakenly interprets Frege as giving reasons for holding the Begriffsschrift to be true:

The basic principles [i.e. the axioms and rules of the Begriffsschrift] gain something from our... recognizing "advantages" of simplicity, sharpness, and the like...Frege is defending not merely the logicity but the truth of his basic proposed principles. (2005, 344)

But there is no evidence that Frege intended these comparisons to convince his reader about the truth of his basic principles. The fact that Frege does not introduce his notation by insisting that its basic principles are self-evident rather implies that he is not worried that his reader will doubt

their truth. He is only worried that his reader will not grasp the value of using his notation.<sup>25</sup>

What the Begriffsschrift “gains” from its advantages is not a justification for us to hold it true, but an underlining of its value for scientific inquiry. By emphasizing its capacity to express gap-free proofs *simply* and to do so for diverse instances of reasoning (i.e. *generally*), Frege gives reason for thinking that his formalism captures *basic* principles of inference, not for holding those inferences justified. Frege employs courses-of-values of functions to represent second-order reasoning in first-order terms.<sup>26</sup> He stresses the *flexibility* afforded users by his formalism since he believes this procedure applicable to any overtly higher-order reasoning.

Secondly, Burge cites Frege’s claims that “fruitfulness is the acid test of concepts” (PW, 33) and that “definitions show their worth by proving fruitful” (FA, §70) as evidence of “pragmatic considerations...[entering] into Frege’s conception of the justification of definitions,” and concludes that “instead of reflection, Frege appeals to mathematical practice...as a way of confirming the worth, and seemingly the correctness, of the definition” (2005, 342). Burge is right that Frege believed that the “fruitfulness” of definitions and concepts was established pragmatically by their value for scientific inquiry (see also FA, §88). He is also right that immediately after giving his definition of the “Number belonging to a concept” as an extension, Frege writes that, since we often think of extensions “as something quite different from numbers,” “That this definition is *correct* will perhaps be hardly evident at first” (FA, §69, my emphasis). Since Frege writes that the truth a definition “contains” is self-evident (CP, 302),

---

<sup>25</sup> Burge is right that Frege takes the advantages of his Begriffsschrift to establish its “logicality.” Burge writes that Frege argues, “that the only way to get a *true* logic is by providing a deeper analysis of judgments and inferential patterns than his Boolean opponents had provided” (2005, 339, my emphasis). But Burge’s use of “true” here is potentially misleading. Frege gives no indication that his own system *truly* or *accurately* codifies ontologically objective logical laws, which Boole’s alternative fails to do. Rather, Frege says that his concept-script is a “fresh approach to the Leibnizian idea of a *calculus ratiocinator*” (PW, 10), and that to the extent that it allows for a deep analysis of judgments and inferential patterns—and is more general than Boole’s language—it deserves to be called “logic.” Frege is evaluating his logic on the basis of its capacity to clarify our grasp of judgments and informally accepted inferences, *not* on the basis of its putative expression of ontologically objective laws.

<sup>26</sup> Frege uses this method when proving his general principle of induction (GGZ1, §117).

Burge takes Frege's account of his definition as "hardly evident at first" to demonstrate his nuanced understanding of "self-evidence." Burge's Frege believed that although *all* of his definitions were self-evident, some were not subjectively obvious, and sought to persuade his reader of their truth pragmatically by demonstrating their fruitfulness.

In the sections following §69, Burge thus sees Frege as advancing a pragmatic justification for *holding* that his definition of number in terms of extensions is true. But there is no evidence that Frege thought that the *truth* of his definition was "confirmed" or justified by the reader's evaluation of its fruitfulness in the way that Burge implies ("...confirming the worth, *and* seemingly the correctness, of the definition" [2005, 341, my emphasis]). The text does not have the structure Burge attributes to it. Instead, having proposed his definition in §68, Frege worries that, given the relative obscurity of extensions of concepts (and in particular their relationship to numbers), his readers will fail to understand it. They may worry, for instance, that if valid inferences about extensions differ from valid inferences about numbers, Frege's proposal will license invalid arithmetical deductions. So in §69, in lieu of a lengthy digression into the technically complex question of what extensions are,<sup>27</sup> Frege aims to quell such anxieties by showing that defining Number in this way does not contradict any of the "basic assertions" about numbers upon which we informally agree. Having bought himself this grain of salt—convincing his reader, that is, that his definition will not *obviously* license invalid deductions—in §70 he turns to the *distinct* question of why we *should* define numbers in this way. Frege proceeds to argue that his definition is fruitful not only because it allows for the definition of Number and for formal proofs of arithmetical truths, but because extensions can be

---

<sup>27</sup> This is not the purpose of *Foundations*, which, after the muted reception of *Begriffsschrift*, Frege intended to write accessibly (COR, 102; 172). He formally defines extensions using Basic Law V in *Grundgesetze* as the courses-of-value of first level functions. In *Foundations*, he simply states "I assume it is known what the extension of a concept is" (§69, n1).

repeatedly and gainfully employed in subsequent deductions involving complex arithmetical structures.<sup>28</sup> This argument confirms the *worth* of the definition, sufficing to recommend its use.

It neither gives, nor is it intended to give, justificatory support for the definition's *truth*. To Frege, alleging a definition fruitful is just to claim it helpful for scientific inquiry.

Thirdly, Burge claims that Frege's recommendation that his reader familiarize himself with the Begriffsschrift notation shows that Frege believed that his reader's "acceptance of [the truth of his basic principles might depend on] detailed mastery of the system" (2005, 344).

Further, Burge claims Frege believed that the laws of his Begriffsschrift might, in principle, be flawed (344). Neither claim is supported by the passages that Burge quotes:

I have drawn together everything that can facilitate a judgment as to whether the chains of inference are cohesive and the buttresses solid. If anyone should find anything defective, he must be able to state precisely where, according to him, the error lies: in the Basic Laws, in the Definitions, in the Rules, or in the application of the Rules at a definite point. (GGZ1, vii)

I myself can estimate to some extent the resistance with which my innovations will be met, because I had first to overcome something similar in myself in order to make them...After one has reached the end [of the book], he may reread the Exposition of the *Begriffsschrift* as a connected whole...In this way, I believe, the suspicion that may at first be aroused by my innovations will gradually be dispelled. The reader will recognize that my *basic principles* [Burge's emphasis] at no point lead to consequences that he is not himself forced to acknowledge as correct. (GGZ1, xi-xii)

Burge claims that in the first passage, "the language clearly suggests that [the] defect might in principle lie not only in the ordering or in the proposed axioms not being logical, but even in proposed basic principles not being true or sound" (344). He claims that the second passage shows that Frege believed a critic might initially *doubt* that a particular axiom or rule of the Begriffsschrift was true, but become convinced of its truth once she had worked through the book for herself and recognized that the basic principle in question does not imply any

---

<sup>28</sup> "[Every element in the] really fruitful definitions in mathematics" is, Frege writes, "intimately, I might almost say organically, connected with the others," allowing for inferences that "cannot be inspected in advance." (FA, §88). He holds up the epsilon-delta definition of continuity as an example. When one proceeds to prove topological theorems, one will continually return to this definition of continuity to argue in terms of a function's limits. Similarly, once Frege proceeds to prove arithmetical theorems, he repeatedly returns to this definition of Number to argue in terms of extensions.

falsehoods (345). But this interpretation neglects the context of these passages. As has been well documented, Frege's *Begriffsschrift* was poorly understood by those of his contemporaries who read it, including its few reviewers, a fact that greatly frustrated Frege (Dummett 1973, xxxi; Gregory Currie 1986, 79). Just prior to the second passage Burge quotes, Frege expresses his opinion that both the *Begriffsschrift* and the *Foundations* have been unfairly ignored (GGZ1, xi). Much of the introduction to the *Grundgesetze* is thus designed to ward off shallow criticism based on confusion.

This casts the passages in a very different light. Frege addresses readers put off by his unfamiliar notation. He encourages such readers to engage his ideas seriously by directing whatever criticisms they might have to the appropriate part of his system, rather than prematurely discounting his innovations on account of their novel presentation. The language of the first passage does not therefore “clearly suggest” that Frege believed that his basic principles might be false (or that his method for codifying logic was fallible). It only shows that he believed *others* might (erroneously) object to his system. Frege's hope is that in the course of making any objection precise—perhaps in *attempting* to argue that one of Frege's basic principles is false—his critic would come to realize that it was *he* who was mistaken. Frege returns to this theme at the end of his introduction: “Anyone who has different [logical] convictions can just try to erect a similar structure upon them, and he will, I believe, come to realize that it does not work, or at least that it does not work so well” (xxvi). If his critic does not succeed in isolating a problem with the *Begriffsschrift* but is nonetheless dissatisfied, Frege challenges him to *show* a better codification of principles of demonstrative inference, rather than letting his vague feelings of disquiet prejudice him against the value of his system.

Likewise, it is not plausible to read the “suspicion” about his “innovations” that Frege attributes to a reader in the second passage as a suspicion that the basic principles are *false*. Rather, it is a wondering that the unfamiliar concepts Frege uses (and the unfamiliar notation he has created) will be of any *use*. This suspicion will be “gradually dispelled” as one realizes that the Begriffsschrift’s concepts are actually precise explications of the vague concepts one has already been using, and that deriving proofs within the formalism clarifies the inferential patterns upon which one has been tacitly relying.

Fourthly, Burge turns to the doubt Frege famously, and presciently, expresses about Basic Law V:

A dispute can break out here, so far as I can see, only with regard to my fundamental law concerning value-ranges (V), which has not yet perhaps been expressly formulated by logicians, although one has it in mind, for example, when speaking of extensions of concepts. I hold it to be purely logical. At any rate the place is hereby indicated where the decision must be made. (GGZ1, vii)

Frege’s attitude in this passage certainly seems puzzling. If he believed his basic principles self-evidently true, what sort of “dispute” can he be thinking of? Burge claims that this remark shows that Frege was “uneasy about [Basic Law V’s] truth” (2005, 348 n21). Since Frege nevertheless advanced it as an axiom, this suggests to Burge that Frege must have thought it a self-evident truth that *he himself* did not (yet) find subjectively obvious. Nonetheless, he thought himself pragmatically justified in holding Basic Law V to be true *because* of the advantages of the Begriffsschrift as a whole.<sup>29</sup>

---

<sup>29</sup> This reading seems to jive with Frege’s comments about Basic Law V after Russell discovered the contradiction: “I have never disguised from myself its lack of the self-evidence that belongs to the other axioms and that must properly be demanded of a logical law...I should gladly have dispensed with this foundation if I had known of any substitute for it. And even now I do not see how arithmetic can be scientifically established...unless we are permitted – at least conditionally – to pass from a concept to its extension” (GGZ2, 253). Frege reports not only that he did not find Basic Law V self-evident (i.e., to Burge, that the axiom was not subjectively obvious), but also that he did not (and does not) see any way to establish logicism without it. On Burge’s interpretation, this suggests that Frege was led to *hold* Basic Law V true (indeed, hold it to be an axiom) despite its lack of self-evidence *because* of his conviction in the truth of logicism, and his inability to see any other way of confirming logicism.

But Burge's description of Frege's state of mind is not the only way to understand the "dispute" that Frege thought might arise about Basic Law V, and thus does not constitute compelling evidence for his interpretation. Jeshion claims that the interpretation she and Burge develop is nonetheless preferable, because it is the only one to allow for a charitable account of Frege's attitude.<sup>30</sup> However, I think it is equally charitable (and plausible) to interpret him as uneasy not about Basic Law V's *truth*, but about its role in clarifying the principles of demonstrative inference. The value of other novel features of his formalism used in the rigorous proofs of uncontroversial truth-functional deductions (e.g. taking truth values to be objects) would become quickly apparent to anyone who familiarized themselves with the notation, and who reflected on the simplicity and elegance of the resulting proofs. But Frege intended Basic Law V to simplify the proofs of mathematical theorems that were being colloquially deduced using a variety of inference patterns.<sup>31</sup> He recognized that any inquirer who doubted one of those inference patterns, or who was unfamiliar with some of them, might fail to grasp the value of Basic Law V.<sup>32</sup> Moreover, he allowed that another logician might discover a simpler way than he to formalize the diverse inferences used by mathematicians. For these reasons, Frege viewed Basic Law V differently from his other basic laws, but did not doubt its truth.

It is also worth emphasizing that Frege presents Russell's contradiction by providing a proof *within* Begriffsschrift that Basic Law V is false: that is, he proves its falsity on the basis of

---

<sup>30</sup> Jeshion (2001, 970). She claims that Philip Kitcher's interpretation makes Frege come out as either insincere or self-deceptive in his attitude toward Basic Law V's fitness as an axiom.

<sup>31</sup> For more on the inferences licensed by Frege's mathematical contemporaries, see Mark Wilson (1995).

<sup>32</sup> Ricketts notes that Frege introduces Basic Law V by informally arguing for the second-order claim that a function *exists* which takes functions to their value ranges, and then stipulates that it can be named by the courses-of-values function (1997b, 200; 210-211). A constructive mathematician might object to introducing a function in this way. Moreover, Frege's contemporaries did not agree upon the legitimate basic assertions about "extensions" (as Burge notes, "the dispute between Frege and the iterative set theorists suggests that there was a fundamental doubt about the viability of the notions (respectively) of set and of extension" (2005, 348, n21). Perhaps Frege worried that, if future mathematicians decided to impose restrictions upon extensions, the definition of Number in terms of extension for which Basic Law V allowed would no longer be fruitful.



his certain grasp of the other axioms in his codification (GGZ2, 256-260). He expresses none of the doubt about his other basic laws that might be reasonably expected if Burge were correct, and Frege thought that his grasp of those laws was as fallible—and now as demonstrably in jeopardy—as his grasp of Basic Law V. Rather, his attitude is consistent with thinking that although he has certainly succeeded in codifying *some* of the principles of inference in the rest of *Begriffsschrift*, he has made a mistake in the isolated case of Basic Law V.

Finally, Burge notes that in §90 of the *Foundations* Frege writes that conclusively establishing logicism would require explicitly formalizing gap-free proofs of all arithmetical truths within the *Begriffsschrift*. To Burge, Frege is suggesting we would be *justified* in holding logicism true through empirical work. Moreover, since logicism is a thesis about the justification of arithmetical truths (not a thesis intended to establish *that* arithmetical truths are true),<sup>33</sup> Burge claims that if empirical work is to establish logicism the *Begriffsschrift* must not only be free of error; its deductions must also “capture an antecedent order of justificational priority” (2005, 345) possessed by thoughts in the third realm. In Burge’s view, Frege posits a link between our scientific inquiry and our knowledge of truths in the third realm.<sup>34</sup> His Frege thinks that we can be *prima facie* confident of our grasp of the ontologically objective truths of logic and mathematics on the pragmatic basis that “pure mathematical practice *works*” (2005, 341, original emphasis).

However, there is no evidence in Frege’s texts that he believed that this “antecedent order of justificational priority” was to be cashed out in Platonist terms, nor that he thought empirical

---

<sup>33</sup> As Frege notes early on in *Foundations*, “After we have convinced ourselves that a boulder is immovable, by trying unsuccessfully to move it, there remains the further question, what is it that supports it so securely?” (§2). Logicism is a thesis about what supports the boulder of mathematics.

<sup>34</sup> In fact, Burge thinks Frege’s most original contribution to Euclideanism is exactly this recognition that pragmatic considerations can play a justificatory role. According to Burge, Frege’s contextualist claim that grasping the self-evidence of logical foundations demands reflection upon, and so familiarity with, the inferential patterns used by scientists both modifies and improves the traditional rationalist claim that such a grasp only requires an understanding of the component concepts appealed to in the expression of logical laws (341).

work could justify holding logicism true. In §90, he writes that empirical work may remove lingering doubts about whether the proofs of some arithmetical theorems rely upon non-logical inferences, but I am not aware of any text where he identifies the removal of doubts with a “justification” for holding true. What an individual finds doubtful is a matter of her psychology, but what is justified, according to Frege, are mind-independent thoughts. Further, it is plausible that the antecedent order of justification that he thought the proofs expressed in Begriffsschrift notation would have to convincingly capture is simply that order upon which working scientists tacitly depend in *practicing* science. A judge familiar with Frege’s notation would then be positioned to evaluate whether it codifies the principles of demonstrative inference without needing to posit a link between her practice and true thoughts in the third realm. She would need only to reflect upon whether the formalism captures the patterns of inference upon which she already relies.

This concludes my argument against Burge’s attribution of B4 to Frege. The previous paragraph is especially suggestive. If Frege did not intend to provide his readers with a pragmatic justification for holding logic true, and if, indeed, his insistence that logic be divorced from psychology makes attributing any notion of “justification for holding true” to him puzzling, then we have reason to doubt Burge’s characterization of Frege’s epistemology of logic. In the next section, I shall examine Weiner’s alternative account of Frege’s project and argue that it better comports with the appeals Frege makes to pragmatic considerations.

### **3. Weiner's Frege: A Logocentric Elucidator**

According to Weiner, Frege is committed to the following theses:

(W1) *Objectivity as Thinker Independence*: Thoughts are “objective” in the sense that they are neither equivalent to nor dependent upon psychological facts about individual judges.<sup>35</sup>

(W2) *Platonism about Thoughts draws attention to key features of inquiry*: Talk of thoughts “existing” in a “third realm” is a rhetorical device used to mark off truth-apart discourse from other uses of language.<sup>36</sup>

(W3) *The Begriffsschrift admits only of elucidation, not justification*: By considering elucidations of the Begriffsschrift, inquirers may come to recognize that they acknowledge the truths that it codifies, but the formalism cannot be justified since any putative attempt to do so would assume its principles and thus be rendered circular.<sup>37</sup>

I have constructed these numbered theses to parallel those Burge defends. Clearly, W3 is incompatible with B3 and B4, but are W1 and W2 opposed to B1 and B2? What is to stop us from adopting all four theses as part of a more complete description of the significance of “objectivity” and “Platonism” for Frege?<sup>38</sup> If this is possible, it may seem that we should interpret the point about which Burge and Weiner are *clearly* opposed—whether Frege offers justifications or only elucidations—as merely a terminological matter.<sup>39</sup> I think that this would be a serious mistake. It conflates distinct interpretations of Frege’s project. In this section I shall justify this claim by showing how, on Weiner’s reading, unlike Burge’s, Frege does not face the

---

<sup>35</sup> Weiner (1990, 170-175; 1995, 364). See also Ricketts (1986).

<sup>36</sup> I will argue below that this is the best way to understand Weiner (1990, 216-218; 1995, 373, 375-376).

<sup>37</sup> Weiner examines Frege’s elucidations in detail in the sixth chapter of her (1990). See also Weiner (2005; 2008).

<sup>38</sup> Both Weiner (1990, 248-250) and Burge (1986, 123-131) develop objections to Dummett’s canonical reading of Frege as using linguistic categories to understand logical categories (where an “object” is whatever a singular term stands for, and other linguistic expressions are defined inductively on the basis of singular terms, corresponding to higher order “concepts” [1973, 54]), which might fuel the thought that their interpretations can be reconciled in opposition to Dummett’s. But while Burge thinks that Frege’s logical categories are developed on the basis of his grasp of the ontological structure of abstract thoughts, Weiner argues that they are developed on the basis of our shared grasp of the norms guiding our inquiry.

<sup>39</sup> According to Greimann, if we define “justifications” as any method that shows us a truth, then Frege’s remarks are justifications, but if we define “justifications” as only those methods whose success relies upon our grasp of an inference, then Frege’s remarks are not justifications (2008, 406). Cf. Richard Heck (2010, 343-344).

epistemological problems of justifying either his demarcation of logic or the truth of his codification to his readers.

Weiner's interpretation belongs to a family originating with Jean van Heijenoort (1967) that emphasizes ways that Frege's conception of logic as a universal language differs from our contemporary conception of logic as the investigation of formal calculi.<sup>40</sup> She reads "judgment," and the practice of judging that characterize our truth-apt scientific investigations, as Frege's starting point (PW, 17-19). In this context, Frege's aim in constructing the *Begriffsschrift* is to create a universal language that will render perspicuous the justification for any judgment, a powerful feature that licenses honoring his formalism as "logic." His conception of logic is thus *logocentric*, originating in the linguistic and conceptual capacities we possess as judges. Since according to Weiner's Frege we have no grasp of justification that is independent of that which finds expression in the *Begriffsschrift*, Burge's attempts to provide justifications *for* the *Begriffsschrift* presuppose what they are meant to demonstrate, and so, fail. So, what does the epistemology of logic come to under Weiner's interpretation of Frege's project?

Since humans must use language as a medium to communicate what they are thinking to each other, practices which involve thinking together (such as scientific inquiry) are necessarily linguistic.<sup>41</sup> But Frege thinks ordinary language is a poor medium for achieving scientific aims. The reference of terms in ordinary language is often ambiguous. Moreover, the inferences that inquirers endorse (and reject) are about the subject matter being considered, not the sentences

---

<sup>40</sup> See especially Ricketts (1985; 1986; 1996), Warren Goldfarb (1982; 2010), James Conant (1991), Michael Kremer (2000), and William Taschek (2008).

<sup>41</sup> Frege thinks it conceivable that telepathic rational beings might exist who have no such need to rely on language when engaged in scientific inquiry (PW, 269). Would such beings have any need for logic as Frege conceives it? On the one hand, Frege thinks *we* can make use of a *Begriffsschrift*—a universal concept *script*—to clarify the justificatory support for (and content of) our judgments, which we can only do imperfectly by relying on ordinary language. Telepathic aliens would have no use for a *Begriffsschrift* in this sense. On the other hand, to the extent that they are making judgments and drawing rational inferences on the basis of their prior judgments, telepathic aliens must be *abiding by* the laws of logic. They could perhaps benefit from a logical education that helped them to explicitly recognize the laws of inference, and so think more clearly.

which they use to discuss that subject matter. There is therefore a danger that some good inferences will be unobvious and some bad inferences will appear legitimate when they are expressed using a natural language with its own grammatical rules. Frege intends his Begriffsschrift as an alternative language within which to conduct science whenever a proof is in question (BGS, v-vii; GGZ1, 3).<sup>42</sup> According to Weiner, the sole principle guiding its design is to be maximally conducive to our scientific aims. Frege does not conceive of logic as grounded upon any substantive prior ontological or epistemological theory. The ontological categories of “object” and “concept” that he invokes have no independent or prior existence—i.e. a Platonic existence—from his logical analysis, his segmentation of the sentences we use to express judgments into object and concept positions.<sup>43</sup> Instead, Weiner’s Frege’s conception of what “logic” is arises from the role to which he hoped his Begriffsschrift could be put.<sup>44</sup> To understand his view of logic and its epistemology, we must unpack his conception of our practice of judgment.

What desiderata are generated by our practice of inquiry for an artificial language specifically designed to aid us? Firstly, and most importantly, Frege thinks judgment aims at a standard that is both *shared* and *objective*. The sense in which the standard is objective is exhibited by our reaction to conflicting utterances when we are in scientific contexts. In some of our other linguistic practices, such as discussions of taste, we interpret conflicting utterances as the *legitimate* expression of a difference of opinion. Features of Anne’s physiology may cause

---

<sup>42</sup> Frege is explicit that “this Begriffsschrift is an aid devised for particular scientific purposes, and should not therefore be condemned because it is no good for others (vi). He does not intend the Begriffsschrift to *replace* ordinary language, which is used for a variety of purposes, many of them non-scientific.

<sup>43</sup> See especially Ricketts (1986, 89-90).

<sup>44</sup> This is part of what Ricketts (1996, 124) and Goldfarb (2010, 72) mean by calling Frege’s conception of logic “retail” not “wholesale.” Frege does not begin with a philosophically developed conception of logic and claim to have elaborated *all* of its laws in his Begriffsschrift. Rather, he begins by articulating *some* laws in the Begriffsschrift he devises, and having reflected upon the Begriffsschrift’s capacity to clarify scientific discourse, recommends that we call it “logic.” (As Ricketts put it to me once in conversation, Frege offers only a “nominal,” not a “real,” definition of the “essence” of logic.)

her to legitimately and sincerely remark that the blue cheese which David has just called delicious is *not* delicious. We think both Anne and David may correctly maintain their opposed positions. But in science, we interpret conflicting utterances as the *illegitimate* expression of a contradiction. If Anne sincerely reports that this cheese is British while David sincerely reports that it is *not* British, but Dutch, both recognize that at least one of their claims ought to be rejected. Anne counts as *asserting* a claim that David *denies*. Our belief that at most one of a pair of conflicting utterances in scientific contexts can be *correct* reveals that scientific inquiry has an objective norm.<sup>45</sup> We have named the “correct” judgments of our practice of scientific inquiry “true,” and we take a central *shared* goal of scientific inquiry to be the discovery of shared truths. It is the standard to which we hold our fellow scientific inquirers responsible, just as inquirers. To Weiner’s Frege, this is our sole understanding of “truth.”<sup>46</sup>

Secondly, scientific inquiry is characterized by *inferential chains*. Judging is not only a matter of endorsing truths in isolation. Many (indeed, most) of the truths we endorse are endorsed *on the basis of* other truths that we have recognized. These chains of reasoning are also subject to a shared standard of correctness in terms of the goal of our activity: truth-preservation.<sup>47</sup> In this way, we credit some inferences as “justified” and reject others as “unjustified.” We also take it to be part of our practice that we may demand of our fellow inquirers that they make clear what justification they have for a particular truth they endorse.

---

<sup>45</sup> Frege distinguishes scientific dialogue from aesthetic dialogue in a number of passages (GGZ1, xviii-xix; T, 69; PW, 233). For discussions of this point, see Ricketts (1986), Conant (1991, 147), and Sanford Shieh (2002, 102).

<sup>46</sup> Frege holds “true” to be indefinable (T, 60), but elucidates it by repeatedly describing judgment as “the recognition of the truth of a thought” (T, 62; PW 139; 185; 197). These descriptions elucidate both “truth” and our practice of “judgment” at the same time. It falls to logical laws to “unfold” the content of the word “true” (PW, 3) by making the principles of inference guiding our practice of judgment explicit.

<sup>47</sup> Often, we hold chains of reasoning to the standard of *likely* truth-preservation (i.e. non-demonstrative inference), rather than *certain* truth-preservation (demonstrative inference). But Frege believes that our grasp of truth is constituted by our grasp of the principles of demonstrative inference (PW, 3). In this sense, our non-demonstrative inferential practice is parasitic on our demonstrative inferential practice.

These features give rise to a number of desiderata for an ideal language for science. Firstly, each expression should have an unambiguous reference, so that discussants can clearly grasp which claims are being discussed. Secondly, if one inspects the sentential expression of an endorsed claim within the notation, its justificatory support should be entirely clear. Thirdly, it should be possible to make clear the justification for *any* endorsed claim within the notation.

It is with these desiderata in mind, according to Weiner, that Frege designed his Begriffsschrift as a universal language for scientific inquiry, a *calculus ratiocinator* (BGS, vi). They guide his choice of technical vocabulary. In his later work, Frege introduces “thoughts” as a term for the objective content that inquirers are trying to express in language, which they grasp through cognitive acts; “judgment” as the recognition by an inquirer that the thought she is currently grasping is true; and “assertion” as the public manifestation of an inquirer’s private judgment (T, 62). There is no conceptual priority to these terms, and so none are defined on the basis of the others (T, 60). Rather, they form an interrelated network that explicates key features of our demonstrative inferential practice. In line with the first and third desiderata, Frege aims for the Begriffsschrift to be universal. It should be possible to construct an unambiguous representation of each candidate for judgment—each thought—in the Begriffsschrift. In line with the second desideratum, Frege aims for the Begriffsschrift to make explicit the inferentially relevant content of each thought.

Frege’s heroic insight was to achieve this second goal by adopting a function/argument analysis (already familiar from mathematics) to evaluate the thoughts expressed by sentences. Inspecting a thought analyzed in this way allowed inquirers to grasp how that thought could be understood to instantiate different sorts of generalizations. This process, as Frege recognized, is at the heart of scientists’ quantificational reasoning.

Frege employs gothic letters as variables to express generality. But the Begriffsschrift sentences in which Frege's variables appear express (as Warren Goldfarb has put it [2010, 69]) *substantive* thoughts, not *schematic* formulae that stand in need of semantic interpretation in order to express thoughts and be either true or false. Since Frege intended the Begriffsschrift to be universally applicable, he took variables to be unrestricted (unlike our current usage, which ties variables to particular domains). However, he introduced different *sorts* of variables correlating to the different *types* of generality that could be expressed if one abstracted away from different parts of a thought.<sup>48</sup>

For example, consider the sentence “Neville is a cat.” Using the function/argument analysis, one can view this as the result of placing the proper name “Neville” in the argument place of the first-level function “\_\_\_\_ is a cat.” One may now construct a general thought by appending a quantifier and a variable ranging over *objects* to the first level *function* one has just identified—for example, the generalization “ $\forall x(x \text{ is a cat})$ .” Or, one can construct a general thought by appending a quantifier and a variable ranging over *functions* to the *object* one has just identified—for example, the generalization “ $\exists \phi(\text{Neville } \phi)$ .” By extension one can also generalize over higher order functions. Since inquirers recognize distinct inferences that rely upon each type of generalization, Frege sorts the variables of Begriffsschrift according to their logical type (GGZ1, §8; §28-29).

Frege articulates Begriffsschrift sentences that, once one has understood the symbols of the notation, are evidently true and not in need of justificatory support, which he calls basic laws (GGZ1, §18; §20). He demonstrates how the expression of the conditional within his two-

---

<sup>48</sup> These “parts” of a thought are themselves determined by the inferences which working scientists employ. For example, take this simple inference:  $7 \times 2 = 14$ ;  $2 = (1+1)$ ; so  $7 \times (1+1) = 14$ . We recognize this as legitimate because of what we call Leibniz's Law, that proper names can be replaced in extensional contexts *salve veritate*. But it is only through our recognition of this inference that we can isolate the category of “proper names” at all, as those terms which appear in simple, non-general identities. For further discussion of this point, see Ricketts (1986, 179).



dimensional notation makes the application of *modus ponens*, which he takes to be a self-evidently truth-preserving inferential rule, entirely perspicuous (§14). In contrast to Burge's Frege, "self-evidence" carries little theoretical weight here. For Weiner's Frege, one will judge a rule of inference or axiom self-evident if it transparently codifies a basic principle upon which one relies when judging. Since Frege never defines self-evidence and Jeshion's explication doesn't perfectly match his texts, I count this deflation of the significance of "self-evident" a point in Weiner's favor. Finally, Frege shows how the justification for various mathematical theorems is clear when those theorems are expressed in the Begriffsschrift. Doing so makes explicit how those theorems are rationally derivable using principles of inference from self-evident basic laws (GGZ1 §49-GGZ2 §179). With the addition of plausible enthymematic premises and definitions, and perhaps further non-general basic laws belonging to particular special sciences, he is confident that his axiomatic approach will allow the justification for any scientific judgment to be perspicuously expressed within his notation.<sup>49</sup> For Weiner, Frege's confidence in the universal applicability of his notation is what licenses calling it "logic" as an honorific: he has created a language that he believes meets the desiderata for an ideal language for science.<sup>50</sup>

We are now positioned to understand why Weiner's attribution of W1 and W2 to Frege cannot be reconciled with Burge's attribution of B1 and B2 to Frege. On her interpretation, Frege's philosophical motives for (and concerns about) his formalism sharply differ from contemporary concerns in the philosophy of logic. In particular, he is not concerned to develop a theory of the workings of natural language, nor to defend a Platonist ontology, nor to advance a

---

<sup>49</sup> See (PW, 21-32; CP, 235). Goldfarb works through an example in (2010, 82).

<sup>50</sup> Frege has an experimental notion of completeness (see especially PW, 38). Although he is confident that the proof of any informally accepted theorem will be expressible within the Begriffsschrift, his worked examples only make this conclusion probable. If he has failed to articulate a principle of inference upon which the proof of some accepted result depends, it would have to be added to his codification (CP, 235).

nuanced Euclidean epistemology. Rather, his sole purpose is to design a tool for use by careful inquirers, a universal language in which proofs will be clear, a project which he calls “logical.” He is not engaged, either, in projects we currently consider to be a central part of the philosophy of logic, such as the formal semantic justification of a notation.

Weiner argues that engaging in some contemporary projects in philosophy of logic would not even have made sense from Frege’s logocentric perspective. Her reading relieves him from the pressure of having failed to give—or having given plainly poor—answers to the epistemological problems that seem pressing on Burge’s interpretation. Since many philosophical statements cannot be justified or even expressed within the *Begriffsschrift*, to Weiner’s Frege much of philosophy was only *elucidatory*. Philosophical remarks do not have the status of scientific assertions that can justify others and which admit of and require justification themselves. Instead, philosophy’s purpose is to help inquirers obtain the perspective necessary to *use* the *Begriffsschrift*, so that they may begin the serious work of science.

According to Weiner, Frege believes that we have no grasp on the notion of truth independently of the practice of judging, and, in particular, of scientists’ private judgments (and public assertions). “Truth” is just our name for the constitutive norm of judgment. Since we lack an antecedent grasp of truth, we cannot *explain* our practice of judging in terms of it. Nor is the logician seeking to grasp “laws of truth” that exist independently of human activity (B1). Although Frege often calls the laws of logic the laws of truth (GGZ1, xvi; PW, 128; 145), he writes—in the aptly named “My Basic Logical Insights”—that the essence of logic “is not contained in the word ‘true’ at all but in the assertoric force with which a sentence is uttered” (PW, 272). Our utterances only carry the “force” necessary to be *assertions* if we advance them as candidates for evaluation by others in our scientific community according to the shared norm

of “truth” that guides our practice of judging. To *be* a judge—one who is able to engage in scientific inquiry—is for one’s utterances to aim at truth, to count as assertions. And *to be able* to aim at truth in one’s utterances and to assert is to be a judge.<sup>51</sup>

Weiner extends this line of thought to deflate Frege’s Platonism. She argues that, for Frege, although “thoughts” are independent of cognitive processes undergone by particular thinkers, they are not independent of the scientific practice in which we rational thinkers engage (W1). Their “objectivity” is not an *ontological* feature to be cashed out by Platonist metaphysics, but a feature of our everyday scientific practice, exhibited by our recognition of a distinction between scientific disputes and non-scientific differences of opinion. This feature of our inferential practice must be respected by the designer of a universal language for science. Quoting Frege’s claim that what is objective is “what is subject to laws, what can be conceived and judged, what is expressible in words” (FA, §26), Weiner argues that Frege’s conception of objectivity is irrevocably tied to what we can express and communicate using our language (1995, 370). In this passage, Frege writes that the North Sea is an objective object whose existence does not depend upon any choice or idea of ours. Burge and Weiner agree that this implies that the North Sea would exist (unrecognized) even if we had never drawn its boundaries. But while Burge draws the strong conclusion that things which exist objectively do so *entirely* independently of human activity (which fuels B1), Weiner concludes only that things which exist objectively do so independently of *whether any human has so far expressed them*, but *not* independently of the language which humans *can* use to express.

---

<sup>51</sup> Ricketts (1996) and Kremer (2000) argue that we should understand Frege’s comment that the laws of logic are an “unfolding of the content of the word ‘true’” (PW, 3), and his break from the Kantian characterization of laws of logic as laws of thought to calling them laws of *truth*, in terms of Frege’s view that truth is the constitutive aim of judgment. The capacity for inferential judgment is the capacity to recognize one truth on the basis of another, and the science of logic studies the laws which govern this capacity.

In Weiner's view, Frege's Platonist remarks about thoughts are only meant to show that "the content, for logical or scientific purposes, of a sentence is neither a physical object (the inscription of the sentence) nor the ideas typically associated with the sentence" (1990, 218). In other words, when we use ordinary language to engage in scientific inquiry together, we do not take ourselves to be voicing privately accessible ideas (such as how food tastes *to us*), but rather publically accessible—and evaluable—claims (W2). Weiner's Frege believes these claims are more clearly expressed in his *Begriffsschrift*.

Weiner also denies that Frege meant to espouse a Platonic *theory* in his philosophical remarks about intersubjective communication (against B2). She gives two reasons. One reason is that Frege says: "I take it as a sure sign of a mistake if logic has need of metaphysics and psychology – sciences that require their own logical first principles" (GGZ1, xix). The practice of judgment (including intersubjective communication) that logic codifies cannot tacitly presuppose a Platonist metaphysics that must be grasped before it is fully justified (1995, 593). Rather, the putative justification of a Platonist metaphysical theory would rely upon logic.

Weiner's second reason is trickier. It concerns the fact that Frege took words he used in his philosophical writings, such as "function" and "concept," to be "defective" for scientific purposes (COR, 141; PW, 177; 239; Weiner 1995, 376). The problem Frege has with such words is that when they are used in sentences they fail to refer to entities of the correct logical category. For example, when the word "concept" is used predicatively ("\_\_\_is a concept"), a meaningful sentence of ordinary language can only be constructed if the unsaturated function is saturated by replacing the blank with a proper name. But since proper names refer to things in the logical category of *objects*, the resulting sentence will always be false (Weiner 1990, 251-252). Frege makes no use of these defective words within the *Begriffsschrift*. There is no need

to do so, since the notation is constructed to make the logical category to which any given expression refers perspicuous by inspection. Nevertheless, he relies upon these words when introducing his notation in ordinary language, lamenting that his “expressions, taken literally, sometimes miss [his] thought” (CP 193). Weiner concludes that Frege only intended these defective words to smooth his reader’s transition from ordinary language to the Begriffsschrift, and that they have no significant role to play once competency with the Begriffsschrift is achieved.<sup>52</sup> But now, if a rigorous statement of Platonism requires mention of functions, it follows that no *theory* of Platonism can be stated within the Begriffsschrift.<sup>53</sup> Since Platonism must stand outside the language of justification for Weiner’s Frege, the purpose of his Platonist remarks can only be elucidatory (W3).

Finally, Weiner denies that Frege’s semantic claims about the Begriffsschrift notation constitute a *justification* for it (against B3). She agrees with Burge that Frege thought the only justification for the truth of the Begriffsschrift axioms, as self-evident truths, was those axioms themselves. Any purported chain of reasoning that seemed to derive the axioms would already

---

<sup>52</sup> Dummett argues that Frege was wrong to think that the thoughts clumsily grasped by “\_\_\_ is a concept” locutions could not be formally expressed within the Begriffsschrift (1973, 216-217), using an analysis that makes use of the somewhat artificial second order predicate “\_\_\_ is something which everything either is or is not.” But Weiner’s point stands despite Dummett’s suggestion. Frege clearly thought it necessary to introduce Begriffsschrift using terms that *he* thought could not be captured within his notation.

<sup>53</sup> To my knowledge, Weiner nowhere makes explicit why a statement of Platonism about thoughts would require us to mention functions, focusing instead on arguing that a general claim like “all functions are atemporal,” which Burge attributes to Frege, cannot be formulated within the Begriffsschrift (1995, 594). One reason might be that, were Frege espousing a Platonist theory in his various remarks about functions, concepts, numbers and thoughts, a *complete* statement of this theory would not only require assertions about thoughts (which *could* be expressed within Begriffsschrift notation)—for example, “every thought is atemporal”—but also assertions about the ontological status of the relationship between thoughts and their components—for example, some general statement describing the relationship between a function (such as “\_\_\_ is atemporal”) and the thoughts in which that function appears (such as “every thought is atemporal”). Although Frege wrote that thoughts, like sentences, had incomplete components (COR, 141-142; 163; 165), it is unclear whether or not he believed that functions were *identical* to the incomplete components of thoughts (see Mark Textor [2009] for an interesting discussion). However, there is presumably *some* connection between functions and thoughts that we should expect a Platonic theory to articulate. But as Frege says, “it is not possible to speak of [incomplete components of thoughts] without turning what is in need of completion into something complete and thus falsifying the real situation” (COR, 141). Weiner’s argument that no general statement about functions can be expressed within the Begriffsschrift then shows that no complete statement of Platonism is expressible within the Begriffsschrift.

be tacitly relying upon them, and would hence be illegitimate. But where Burge finds in Frege's semantic claims a *justification* for the soundness of *modus ponens* as an operation on expressions of his logical notation, Weiner finds only elucidatory instruction designed to introduce readers to the application of a self-evidently valid rule in an unfamiliar language (2005, 339).<sup>54</sup>

Unlike Burge, Weiner does not think Frege's semantic claims are hesitant steps towards the sorts of justifications that have since become familiar with the development of model theory and proof theory. She thinks that the notion of employing a truth predicate applying only to entities in a *particular* domain to construct arguments meant to offer logic justificatory support is wholly foreign to Frege's *universalist* conception of logic (2005, 346-347). Any such appeal would suggest that logical laws hold *on the basis* of some restricted domain—such as the *linguistic* domain in the case of “semantic” justifications—which would have its own laws. This would undermine logic's (and, were Frege's logicist project to have succeeded, arithmetic's) maximal generality.<sup>55</sup> Weiner's Frege believes that “truth” is just as defective a word as “concept.” Although it is crucial for elucidating the Begriffsschrift (in particular, for naming what its rules of inference preserve [1990, 260]), it is likewise apt to confuse since its ascription adds nothing to the sense of the sentences in which it appears, and it is not a *predicate* of thoughts (2005, 350-351; CP, 354; T, 59-60). Once again, Frege never uses the defective word “true” in his Begriffsschrift. It occurs only in transitional remarks designed to clarify the Begriffsschrift notation (1990, 261). For Weiner, Frege's elucidatory remarks are intended only to help his reader better grasp his ideas so that they may make use of his innovations in

---

<sup>54</sup> The details of this argument would take us far afield here. In discussing *modus ponens*, Frege uses the expression “is the True,” which Burge takes to be a truth predicate allowing Frege to construct soundness arguments. Weiner argues that this expression is not, nor can it be turned into, a truth predicate holding of all true thoughts; rather, it is a predicate holding of a single object, the True (see Weiner 2005, section 3). Instead, she reads Frege as trying to bring out the self-evidence of *modus ponens* to his reader by giving an informal explanation of *modus tollens*.

<sup>55</sup> Weiner (2005, 347) develops a reading of Frege (CN, 336-340) that makes analogous points about how foreign the idea of justifying logic by appeal to a special science that governed the restricted realm of *thoughts* would be to him.

subsequent inquiry: “the aim of elucidation is to have a particular psychological effect. The aim of gapless proofs from primitive laws of truth is to provide justification” (2008, 434).

The striking differences between Weiner’s and Burge’s interpretations of Frege’s project show why it would be a mistake to interpret her focus on “elucidation” as a mere terminological variant of his “justification.” For Weiner’s Frege, the Begriffsschrift is intended to be the universal language of justification. Formalizing a judgment within the Begriffsschrift makes its justification plain. But the truth of the basic principles of Begriffsschrift cannot be justified by appeal to other truths. I am persuaded by Weiner that Frege did not think his semantic claims constituted a bottom-up justification of the Begriffsschrift’s basic inference rules (against B3). My argument in section two against the pragmatic top-down justification that Burge attributes to Frege (B4) complements Weiner’s reading. As Frege conceives of his project, there is neither a need to provide nor a possibility of providing a *justification* to his readers for why they should hold the Begriffsschrift true. There is only a need to elucidate its self-evidence when teaching it to them, and to appeal to pragmatic considerations to drive home to them its usefulness for their inquiries.

Moreover, on the reading that I am recommending, Frege neither needs nor gives a precise demarcation of logic. Since he does not appeal to “self-evidence” as a general criterion by which one may determine whether a given truth is logical, there is no worry that his mention of self-evidence marks a fall into psychologism. These claims may seem puzzling. One might think that if Frege’s logicist claim is to have any bite, he ought to have a substantive account of what *makes* a truth a logical truth. Furthermore, one might think that to evaluate the success of his Begriffsschrift as a codification of logic, one needs first to have a substantive account of what

logic is in order to compare the Begriffsschrift to it. But both of these objections mistake Frege's project. Let us take them in turn.

Frege created the Begriffsschrift to serve as a precise language in which to establish logicism. He proposes that scientists make use of this language when they need to better understand a disputed proof (BGS, v-vi), a feature that gained prominence in his thought after the discovery of the contradiction shattered his faith in logicism.<sup>56</sup> Calling the basic laws of his language "self-evident" is just to say that a competent judge who understands them will recognize them as true and basic. His logicist claim is that it is possible within the Begriffsschrift to construct a gap-free proof for any arithmetical theorem using only the language's basic principles. So Frege does not need a general criterion by which to determine whether a given truth is a logical truth. He only needs to show that a gap-free proof for any arithmetical theorem can be derived on the basis of the basic laws that he has identified.

Secondly, Frege thinks that a reader's evaluation of the Begriffsschrift is a matter of her grasping its usefulness, not a matter of judging whether it accurately demarcates which truths are "logical." In Frege's view, neither he nor his reader have an independent conception of "logic" against which the Begriffsschrift can be compared. Richard Heck has plausibly argued that Frege knew that there were undecidable sentences (2010, 348).<sup>57</sup> So he could not, for example, have thought that all truths expressible using only the topic-universal vocabulary of the Begriffsschrift (which is one account of what makes a truth "logical") were derivable within the Begriffsschrift. Instead, Frege thought the only standard against which to evaluate a purported

---

<sup>56</sup> In "What may I regard as the result of my work?" Frege writes, "It is almost all bound up with the concept-script" (PW, 184).

<sup>57</sup> For example, it is either true that all objects are value-ranges, or that there is an object which is not a value-range; but neither claim is derivable within the Begriffsschrift. See also (GGZ1, §10), where Frege states that the claim that The True is the extension of the concept under which only The True falls is not derivable, but nevertheless stipulates that it is true.



codification of logic is experimental completeness (PW, 38). He is extremely confident that the Begriffsschrift suffices for his immediate purposes of establishing logicism. He thinks a gap-free proof of any accepted arithmetical theorem can be derived within his notation. He is also fairly confident that, with suitable expansions of its vocabulary, his notation can be employed to express the proof of any scientific claim (BGS, vi). But, he acknowledges that other rules of inference might have to be introduced into the language in order give formal gap-free proofs for some scientific truths (PW, 363). This shows again that Frege did not intend his Begriffsschrift to be the final word on logical truth, and that he did not intend its evaluation to be a matter of judging whether *every* truth worth calling “logical” had been codified by his Begriffsschrift.

However, there is a question closely related to the demarcation problem which Frege does engage: what accounts for the special status of logic with respect to our investigations? I have argued that Frege provides no account of how logical truths may be distinguished from non-logical truths. Maximal generality is a necessary, not a sufficient, condition for logicity. So, why does logic normatively govern all rational inquiry? Understanding Frege’s subtle answer to this question is the goal of the next section. I will argue that Frege’s discussion of logical aliens exposes his *constitutive* view of the relationship between logic and cognition. Further, I will argue that Burge’s reading cannot accommodate Frege’s constitutive view.

#### **4. Madness and Logical Aliens**

One of Frege’s aims in the preface of *Grundgesetze* is to combat a psychologistic conception of logic.<sup>58</sup> He insists that logicians who fail to sharply distinguish their work from psychology—“psychological logicians”—mischaracterize both the *normative force of* and the

---

<sup>58</sup> Frege blames the prevailing psychological attitude for the critical neglect of his logicist program (GGZ1, xi). The empirical investigations of psychological logicians into how *humans* think serve, Frege believes, to make logic seem irrelevant to the mathematical and non-psychological question of what constitutes the foundations of arithmetic.

*class of judges obligated by* logical prescriptions. If, as psychological logicians believe, logical laws are empirically discoverable generalizations of the way that humans think, then they issue in conditional prescriptions relevant for *human* judges (and those who wish to judge *humanly*). Roughly: *if* one wishes to judge as the majority of humans do, *then* one must abide by the laws of logic. Such laws have a normative force over human judgment much as grammatical laws have a normative force over a human language: *if* one wishes to speak English as the majority of English speakers think it is correct to do, *then* one must abide by the laws of English grammar. Frege objects that this position fails to adequately distinguish the real topic of logic—the laws of what is *objectively true*—from what most humans *hold true*.<sup>59</sup> The grammatical laws of a language change over time with the shifting intuitions of its speakers. Similarly, were our brain chemistry to undergo a radical change so that most humans began to freely contradict themselves, then under psychologism the “logical” law of non-contradiction would have to be replaced by a law of contradiction (c.f. PW, 148).

In contrast to the psychological logician, Frege insists that logical laws do not ground *conditional* obligations when judging to comport with a mean, a mean that is based on contingent generalizations, but rather ground *unconditional* obligations upon all judges to meet the objective, unchanging standard of truth: “the laws of [logic]...are boundary stones set in an eternal foundation, which our thought can overflow, but never displace” (GGZ1, xvi). Frege holds that judgment is the advance from grasping a thought to recognizing it true (T, 62). To judge *correctly* is thus to recognize of truths that they are true. Insofar as scientific laws express

---

<sup>59</sup> In Frege’s view, although the connection between logic and objective truth is usually “admitted at the outset” of a logical inquiry, the human, psychological matter of *holding* true often wends its way back into subsequent discussion (GGZ1, xv). He is not only opposed to logicians who explicitly commit themselves to psychologism, (which is fortunate, since few would accept Frege’s characterization of psychologism, and he might accordingly be thought to have erected a straw man), but also to any logician who fails to resist the “corrupting influence” of psychology in her work.

general truths, judges are obliged to judge in accordance with them. Such prescriptions do not merely obligate *human* judges, or judges who wish to judge *humanly*, but *all* judges. Among scientific laws, logical laws deserve the name “laws of thought” not in virtue of a special relationship to empirical psychology, but because they have a maximally general scope in rational inquiry. Logical laws hold universally for every act of judgment, not merely for judgments in a particular special science, such as geometry or physics (GGZ1, xv).

Nevertheless, Frege’s subsequent discussion of logical aliens (beings who deny a logical law) reveals that he thinks there is a crucial aspect of logical laws which is *not* shared by other scientific laws. He thinks that they enjoy a privileged, and what I shall call *constitutive*, relationship to cognition.<sup>60</sup> I think that Frege’s recognition of this relationship explains why he writes not that logical aliens are merely *wrong* for denying a logical law, but that they are *mad*. (Notice that we experience no temptation to call a person who denies a law of physics mad, rather than wrong or, perhaps, foolish). On the basis of my articulation, I will argue that Burge’s pragmatic Euclidean interpretation of Frege cannot satisfactorily accommodate his constitutive view. Furthermore, and in line with Weiner’s logocentric reading, Frege’s constitutive view explains why he was content to rely upon elucidating the self-evidence of the Begriffsschrift to other rational inquirers, since it exposes that he believed such inquirers tacitly endorse logic by engaging in demonstrative reasoning.

Those critics who discuss Frege’s logical aliens rarely note that he identifies two distinct types of alien, and none, in my view, adequately attend to this fact’s significance. A disastrous consequence of underdescribing Frege’s text here is that it makes his constitutive view seem

---

<sup>60</sup> The word “constitutive” is used by a number of commentators when discussing Frege’s conception of logic, including James Conant (1991, 148), Burge (1992, 316), John MacFarlane (2002, 37), and William Taschek (2008, 384). But when we begin to unpack these scholars’ substantive disputes, it becomes clear that they do not share the same ambition in attributing a constitutive view to Frege. One goal of this section is to provide a sharp articulation of the constitutive view to which I think Frege’s texts show him to be committed.

unsophisticated, which in turn encourages (on the basis of charity) thinking it peripheral to his conception of logic. But I think that, on the contrary, the distinction Frege makes between alien species (and, in particular, the fact that Frege only calls members of the *second* alien species mad) demonstrates the subtlety and contemporary interest of Frege's position. Moreover, Frege's discussion of logical aliens shows that he thinks any viable conception of logic must appreciate its constitutive role in inquiry (which, notably, he thinks that a psychological conception fails to do). If I succeed in arguing that interpreting Frege as a pragmatic Euclidean is incompatible with attributing a constitutive view to him, it will be a decisive blow against Burge's reading.

Frege first instructs us to imagine discovering beings who make judgments which contradict one of our logical laws. This is the *first* species of logical aliens. Suppose the law in question is the simplified version of Basic Law I:

[1] 

In modern notation, this is  $(A \rightarrow A)$ .<sup>61</sup> The contradictory judgment made by this first species of logical alien will therefore be, for some thought *B*:

[2] 

---

<sup>61</sup> Basic Law I is actually:



This denotes the judgment, for thoughts *A* and *B*, “it is not the case that *A* is denied, *B* is affirmed and *A* is affirmed” (GGZ1, §18), which we would write in modern notation as “ $(A \rightarrow (B \rightarrow A))$ ”. However, in the case where *A* and *B* denote the same thought, by Frege's rule of the amalgamation of identical subcomponents, this axiom entails the simpler formula [1] above, and which he writes “is a particular case of [the first axiom, and] which will be understood together with [it] without explicit notice” (GGZ1, §18). Here, to make the exposition of Frege's argument as clear as possible, I have chosen to focus on the simpler case, but the argument would go through with the more general axiom—or, indeed, any of Frege's other basic laws.

In modern notation, this is  $\sim(B \rightarrow B)$ , which is equivalent to  $\sim B \& B$ . Perhaps we witness the aliens concluding  $B$  as the result of Monday's investigation, and denying  $B$  as the result of Tuesday's investigation. Balking, we accuse them of contradiction, and show them that they have committed themselves to [2]; but instead of conceding their error and withdrawing one of their judgments, they hold steadfastly to both.

What should we say about such creatures? To the psychological logician, logical laws are abstracted from instances of human thinking. Our current sample of humans do not typically contradict themselves—and, once they have been alerted to the fact in the rare cases where they do, they quickly withdraw their contradictory assertion—which justifies taking [1] to be a logical law. But this tells us nothing about what other beings do. All we may conclude from the alien's judgment [2], then, is that [1] is not a logical law for them.

Frege disagrees. He writes that we ought to conclude that the first species of aliens “do not as we do immediately recognize certain truths, but have to rely perhaps on the longer path of induction” (GGZ1, xvi). His mention of induction may initially seem to discredit my choice of Basic Law, given that [1] has no quantifiers. But Frege understands *all* of his axioms to express general truths:  $A$  in [1] ranges over all objects.<sup>62</sup> Whereas *we* are able to “immediately” recognize that [1] is true, Frege suggests that the aliens' judgment [2] is evidence that they have not (yet) done so. Attributing inductive reasoning to the aliens intelligibly explains this delay. The aliens will only endorse [1] as a *general* truth after explicitly considering a sufficient number of thoughts for them to consider it inductively established. If they have yet to carry out this empirical work, they may reasonably (though, as we know, wrongly) judge [2]. Doubtless, they will soon get into difficulties thinking about claims related to  $B$ . In time, we can show the

---

<sup>62</sup> The horizontal function takes all those objects which are not the True to the False, and takes the True to itself.

inductive aliens that they have ample evidence for [1], and that violating it has consistently gotten them into trouble. We are entitled to judge these aliens *cautious*, but not *crazy*.

I think we can make analogous points about various logically limited inquirers, such as those of us who are prone to logical mistakes. Frege finds it quite intelligible that an inquirer may sincerely make a judgment that she fails to realize contradicts a logical law, as when “our thought...overflow[s]” the laws of logic (GGZ1, xvi). Upon encountering such a person, we may try to explain her mistake *to* her through logical instruction and endeavor to improve her reasoning ability. Extremely limited inquirers may need sustained instruction, but we are only entitled to judge those who have chronic difficulty avoiding inferential errors *stupid*, not *crazy*.<sup>63</sup> Similarly, Frege can make sense of *bold* reasoners who quickly endorse inferences from which we temporarily refrain, *locally eccentric* and *locally foolish* reasoners who systematically endorse invalid inferences in esoteric or difficult fields respectively (while making few such errors elsewhere), and *formally inept* reasoners who fail introductory logic classes but obtain college degrees.<sup>64</sup>

However, Frege asks us to imagine a *second* species of alien that he finds more problematic: “what if beings were even found whose laws of thought directly contradicted our own and therefore frequently led to contrary results in practice as well?” (GGZ1, xvi). It is worth getting clearer here about the case Frege has in mind. Members of the first alien species intelligibly have a limited, inductive form of *our* reasoning. This makes them capable of failing to recognize their logical errors *as* errors, even after we have brought them to their attention. In

---

<sup>63</sup> It is worth noting that psychological logicians *could* likewise interpret these as cases of cautiousness, foolishness, etc. But Frege differs from them in thinking that so interpreting the aliens (rather than judging that, for example, [1] is not a logical law for them) is the *only* intelligible option.

<sup>64</sup> Had Frege been aware of subsequent formal developments, I believe he could likewise have made sense of those *competing logicians* (i.e. intuitionists subscribing to Heyting’s calculus) whose alternate (but not contradictory) inferential practices could be coherently explained as issuing from a guiding principle (such as constructivism). Another example is Niko Kolodny and John MacFarlane’s (2010) argument for a restriction on using *modus ponens* in cases of conditional obligation.

contrast, we are asked to imagine that members of this second species reason in accordance with a law that contradicts one of our own. When we try to draw their attention to their errors, they insist it is *we* who are mistaken. Moreover, the entire alien community appears consistent in their anti-logical verbalizations. They exhibit none of the corrective behavior to which a member of our community would be subjected if they were to judge against a logical law. This community agreement grounds a logically alien “practice” of inquiry, the putative conclusions of which stand in contradiction to our own judgments.

Let us focus once again upon Basic Law I. Suppose its negation is among the second alien species’ laws of thought. These aliens not only contradict themselves, but appear to do so deliberately, stringing together sequences of contradictory judgments in ways that seem meaningful to the other aliens, but which strike us as patently absurd. They not only endorse [2] in a particular case, but act as though it is axiomatic *for every* judgeable content *B*. Let us stipulate that they refuse all of our attempts to correct them (perhaps *they* sometimes seem to be trying to “correct” *us*) and thus call their overt and reflective denial of our logical law [1] “resolute.” Is this second case intelligible?

The psychological logician, Frege suggests, would take the resolute alien’s behavior in her stride. She will conclude that [2] *is* a logical law that is valid for the aliens, while our logical law [1] is valid for us. Frege is damning about this response, because it exposes that the logical laws recognized by psychological logicians are “valid” only relative to a population. Truth, in contrast, is a standard that is “objective and independent of those who judge” (GGZ1, xvii). Truth holds fast for *every* judge. This is the standard which, in Frege’s view, is logic’s proper aim.<sup>65</sup> Without such a shared standard, there is no principled way to distinguish the resolution-

---

<sup>65</sup> A number of commentators have pointed out that nothing Frege says here convincingly *refutes* psychologism (Ricketts 1986b, 69; Shieh 2002, 102). The psychological logician may simply insist, for instance, that a “truth”

demanding contradictory attitudes that characterize scientific disputes from the merely differing opinions that underlie disputes about taste. If you deny my assertion that this powder is copper sulphate, it would always remain an open possibility that our laws of thought differ, much as our taste buds do, and that we have *both* judged correctly in accordance with our distinct laws of thought. This powder might be copper sulphate for me, but not for you. Instead, doing justice to the distinction we commonly recognize between contradictory judgments and differing opinions requires accepting that “anyone who has once recognized a [logical law] has thereby also recognized a law that prescribes how judgments should be made, wherever, whenever, and by whomever they may be made” (GGZ1, xvii).<sup>66</sup> In Frege’s view, our response to members of the second alien species who purport to reason in a way that contradicts our logical laws should be unyielding: “I would say: here we have a hitherto unknown kind of madness” (GGZ1, xvi).

How should we understand this diagnosis? James Conant suggests that Frege chooses “madness” as a way of meeting the psychological logician halfway, that “the closest Frege can come to...[understanding] an antithetical form of reason (deeply illogical thought) is the philosophically innocuous idea of a degenerate form of reason (merely lunatic thought)” (1991, 149). On Conant’s reading, the madness exhibited by logical aliens is of a “hitherto unknown kind” because we have yet to encounter beings who resolutely deny, to stay with our example, Basic Law I. Were we to do so, we could straightforwardly classify them on a continuum with other lunatic thinkers, whose powers of reason are similarly impaired. A perfect capacity for

---

about the world just is what *counts for us* as a truth about the world, which in turn depends (in some way to be filled out) upon the constitution of our minds, and which she has set about clarifying in her inquiry into logic. In later chapters, I shall return to this theme in exploring how the rise of behaviorism in psychology complicates the division between Frege’s conception of logic and that of his psychological opponents.

<sup>66</sup> The committed psychological logician will not find this argument compelling. If she thinks she has empirically discovered logical laws that hold for all members of a given population, then her logic will still be a useful standard for arbitration between members of that population who are engaged in inquiry together. She will insist that to two beings who *share* laws of thought (including, presumably, all humans), this powder is either copper sulphate for both, or for neither. This attitude amounts to a refusal to draw any sharp distinction between contradictory judgments of truth and differing opinions of taste.



logical inference is necessary for perfect thought, but a being displaying *some* capacity for logical inference is intelligible as having a degenerate, “mad” form of thought.

However, I think that Conant’s account fails to do justice to the sophistication of Frege’s thinking. In my view, Frege thinks that the *resolve* of the second alien species ought to *undermine* our initial judgment of their capacity for reason. They do not exhibit a deformed form of reason. They lack *any* form of reason. We are judging beings. They are non-judging organisms. Whatever cognitive activity their physiology allows for does not count as the grasping of thoughts, whatever verbalizations they make do not count as assertions that particular thoughts are true, and whatever series of noises they eject at each other do not count as chains of reasoning.

So calling resolute aliens “mad” is only the first part of Frege’s story. Having initially supposed them rational (perhaps on the basis of our apparently successful early attempts to communicate), their subsequent behavior refutes our charitable hypothesis that they are fellow judges. So long as we abide by our earlier belief and think that we *were* communicating with the aliens, we will be inclined to judge the aliens “mad” on the basis of their now baffling behavior; but as we soberly reflect upon the situation, we will recognize that we were no more communicating with them than we do with our pets.<sup>67</sup>

Conant’s reading results from conflating the two alien species that Frege distinguishes. Frege certainly thinks that we may discover beings who exhibit a limited capacity for logical inference. But, such beings belong to the *first* category of logical alien, not the *second* category

---

<sup>67</sup> We might compare Descartes’ claim that we can infer from the speech of madmen, but not from the squawks of parrots, that thoughts are being grasped, because although the speech of the mad “does not follow reason” it is “relevant to particular topics” (1998, 43). But our judgment of whether or not a being meets Descartes’ standard of her speech being “relevant to particular topics” is really a function of our attempts to communicate with her. The person whose pronouncements we come to judge to have no bearing upon our conversation is—like the parrot—not grasping thoughts, and so is non-rational. (Much as non-human, non-rational animals can be said to “think” about their environment if we interpret their behavior as a series of appropriate responses to the world, Frege may admit that resolute aliens who appear able to negotiate their environment can be said to “think,” if metaphorically.)

whom we are forced to judge mad. If we take the time to instruct a logically limited judge, her reasoning may or may not improve. But the crucial point is that *all* of these impaired reasoners exhibit behavior which is sufficient to make them intelligible as fellow judges who are attempting—just as we are—to recognize true thoughts. All of these judges, *simply as judges*, can be brought not only to endorse particular formulations of the logical laws, but to recognize that the logical laws express the norms of the very practice in which they, and we, are engaged.

Distinguishing the two species of logical aliens serves to make Frege's constitutive view far more plausible than it may seem from other recent accounts. William Taschek, for example, writes that Frege believes all judges must "acknowledge" the laws of logic, but declines to "explicate with any precision what exactly this acknowledgement consists in," on the grounds that doing so would be "very tricky" (2008, 384). But Taschek does note that acknowledgment must "involve one's possessing a capacity to recognize—when being sincere and reflective, and possibly with appropriate prompting—logical mistakes in one's own judgmental and inferential practice and that of others" (384). We can now deepen and clarify this remark. The "appropriate prompting" is not merely a flat-footed demand that one's interlocutor must just *see* her mistake once it is pointed out *because* a particular logical law is self-evidently true. Rather, Frege recognizes that sensitive and creative attempts may be necessary to understand the root of one's interlocutor's mistake, from a variety of intelligible possibilities. His constitutive view only denies cognition to those beings who, like the second alien species, systematically and unrepentantly violate our inferential patterns, who repeatedly contradict our conclusions, and who thereby prevent us from thinking them members of our inferential practice. In real cases, we may of course vacillate in our judgment as to whether the quirky aliens we have just met are deformed and like the first alien species, or resolute and like the second. But this vacillation in

our judgment of others' reasoning capacities does not affect the point that, for Frege, there is a sharp conceptual distinction between *judges* (however deformed) and *non-judges*.<sup>68</sup>

The constitutive role that Frege awards logic makes ineliminable reference to *our* practice of inquiry as *the* practice of rational inquiry:

**CON:           The ability to reason demonstratively as a member of our inferential practice is a constitutive, necessary condition for being a judge.**

Frege writes that the

impossibility of our rejecting [the law of identity] does not prevent us from supposing that there are beings who do reject it; but it does prevent us from supposing that these beings are right in doing so; it also prevents us from doubting whether we or they are right. (GGZ1, xvii)

His point is that reflecting on logic's constitutive relation to judgment tells us nothing about the kinds of things other beings might *seem* to claim about our logic. Other beings might utter sentences which appear to be rejections of a logical law. Nevertheless, Frege thinks that once we have understood the close relationship between logic and judgment, we must judge such beings as *wrong* for doing so. The details of our judgment will differ depending upon the case. If we judge that the aliens are like members of the first species, fallible judges whom we can correct, we will judge that they mistakenly *hold true* the negation of the logical law. If we instead judge that they are like members of the second species, we will judge that they lack the cognitive sophistication to count as judging or holding true *any* thought whatsoever.

Beings can fail to be perfect judges by exhibiting logically limited inferential behavior; indeed, few of us are perfect judges. To count as exhibiting our inferential behavior is thus to possess an inclination to generally (though not always) make valid inferences, to be disposed

---

<sup>68</sup> Similarly, nothing about Frege's constitutive view rules out the possibility of devious aliens. Aliens who initially appear to be like the second species, but who later rescind their deviant activity—perhaps revealing that they had been joking with us all along—may well be forgiven and welcomed back into the fold of judges. Nevertheless, we have a *prima facie* reason, subject to future disconfirmation, to doubt that a being who resolutely violates the inferential pattern codified by one of our laws of logic is a thinker.

(though not always) to endorse instances of the laws of logic, and to sometimes (though not always) correct fellow judges who fail to meet these standards of inquiry on particular occasions.

One might object that this constitutive view cannot be Frege's because of his famous conclusion to the discussion of logical aliens:

Leaving aside logic, one can say: we are forced to make judgments by our nature and external circumstances, and if we make judgments, we cannot reject this law—of identity, for example; we must recognize it if we are not to throw our thought into confusion and in the end renounce judgment altogether. I do not wish to either dispute or endorse this view and only remark that what we have here is not a logical implication. (GGZ1, xvii)

If Frege subscribes to **CON**, why does he not endorse the view that “we must recognize [the law of identity] if we are not to throw our thought into confusion and in the end renounce judgment altogether”? Barbara Massey goes so far as to suggest that Frege only thinks that our logical laws are so central to *our* thinking that were we to meet logical aliens, we would dismiss them as mad, not that the aliens *would* lack cognition (1991, 110).

But, I think that this reading misses that the context of this passage is to reject psychologism. Having just noted the futility of trying to justify a logical law (since all inferential justification *relies* upon logic and so the purported justification would be circular—resulting only in questions of how further logical laws are to be justified), in this passage Frege is considering (and distancing himself from) the way a psychological logician might attempt to justify a logical law. He speculates that the psychological logician might think two results of her empirical investigation pertinent: firstly, *humans* must make judgments in order to survive; secondly, the laws of logic are basic to how *we* judge. Taken together, Frege supposes that the psychological logician will believe these results give humans sufficient reason—pragmatic perhaps, but sufficient nonetheless—to judge the laws of logic true. But Frege once again objects to the psychological logician's emphasis on human psychology—on the fact that *humans* must make judgments, on how *we* judge. He refrains from endorsing or disputing these two

claims, *conceived as psychological results to be empirically established*. After all, Frege has done no empirical research into human thinking, and believes such work (were it to be completed) strictly irrelevant to logic as he conceives of it. This is his point: the psychological logicians' purported justification "is not a logical implication," i.e., cannot justify that the logical laws *are true*, but is merely a psychological implication, i.e., can only justify why *humans* may *hold* logical laws true. His constitutive view means that he is indeed committed to the claim that "we must recognize [the law of identity] if we are not to throw our thought into confusion." But he believes that this is true *because* the laws of logic have a constitutive role in judgment, and does not *ground* that claim; hence it cannot *justify* our judging logical laws true.

Why does Frege believe that logic has this special, constitutive relationship to cognition? What distinguishes resolute violations of logical inferential patterns from resolute violations of inferential patterns in a special science, and what justifies a diagnosis of madness in the former case as opposed to mere foolishness or quackery in the latter? I think that we can reconstruct the answer from a passage in *Foundations of Arithmetic*:

For the purposes of conceptual thought we can always assume the contrary of some one or other of the geometrical axioms, without involving ourselves in any self-contradictions when we proceed to our deductions, despite the conflict between our assumptions and our intuition...Can the same be said of the fundamental propositions of the science of number? Here, we have only to try denying any one of them, and complete confusion ensues. *Even to think at all seems no longer possible*. The basis of arithmetic lies deeper, it seems, than that of any of the empirical sciences, and even than that of geometry. The truths of arithmetic govern all that is numerable. This is the widest domain of all; for to it belongs not only the actual, not only the intuitable, but everything thinkable. Should not the laws of number, then, be very intimately connected with the laws of thought? (FA, §14, my emphasis)

This passage is further evidence of Frege's constitutive view.<sup>69</sup> He infers that since denying a "fundamental proposition" of arithmetic throws us into such confusion that "to think at all seems

---

<sup>69</sup> I thus disagree with Øystein Linnebo's suggestion that Frege was given to immature hyperbole in the *Foundations* (2003, 3). In his mature work, Linnebo suggests that Frege had learned to disavow any attempt to tie logic to what seemed unthinkable to humans as irredeemably psychological. But Frege's rejection of the psychological logicians' putative *justification* for logical laws does not tell against the claim that he held a constitutive view about logic's relation to cognition throughout his career.

no longer possible,” arithmetic is “very intimately connected” with the “laws of thought,” i.e. the logical laws. So, he is claiming that one can no more deny an arithmetical law than one can deny a logical law without thereby abnegating one’s status as a thinker, from which it immediately follows that Frege thinks denying a logical law abnegates one’s status as a thinker.<sup>70</sup> But this passage also reveals that the reason for his constitutive view is tied up with the maximal generality of logico-arithmetical laws. This justification needs to be unpacked.

To see how Frege is viewing these matters, let us first consider the geometrical case he raises. Frege, following Kant, views the axioms of Euclidean geometry as expressing the limits of our intuition of space. In other words, he thinks that it is a fact about the cognitive constitution of humans that all of our spatial experiences can be represented within a Euclidean framework. Even our most imaginative fantasies (if they involve spatially extended entities) are subject to Euclidean geometry. Nevertheless, Frege’s mathematical contemporaries had demonstrated that they could “leave the ground of intuition entirely behind” (FA, §14), and meaningfully reason about *non*-Euclidean geometries. To Frege, such geometers are intelligibly engaging in non-contradictory hypothetical reasoning because their activity can be represented within the Begriffsschrift. They endorse conditionals whose antecedents are the conjunction of Euclid’s axioms minus the parallel postulate, and whose consequents are (contrary to intuitable fact) theorems of non-Euclidean spaces. Their chains of reasoning about these conditionals are logically certified inferences. Moreover, enough familiar Euclidean theorems are provable on the basis of this restricted conjunction of axioms for the resulting structure to be intelligibly recognizable *as* a “space.”<sup>71</sup>

---

<sup>70</sup> It should be clear that the failure of Frege’s logicist project does not rob his constitutive view of contemporary application. One can hold that logical laws are constitutive of thought, even if arithmetical laws are not.

<sup>71</sup> In 1854, Bernhard Riemann had modeled a non-Euclidean space by interpreting lines drawn on the curved surface of a sphere as straight. For Frege, although our intuition of this model (like all of our intuition) is Euclidean (as we

Frege believes that the same is not true for arithmetic. One reason might be that, unlike the geometrical case, the set of arithmetical axioms is too paltry to withstand a deletion and still ground counterfactual results that intelligibly concern an *arithmetic* as opposed to a mere *series*.<sup>72</sup> But I think that a different reason explains why Frege identifies “self-contradiction” as the distinctive problem with any attempt to reason having hypothetically denied an arithmetical proposition.

Suppose that the basic arithmetical proposition we deny is  $1+2=3$ , and that we know from scientific investigation that there is exactly one F, exactly two Gs, and that no F is a G. Having denied  $1+2=3$ , we conclude that there are not three F-or-Gs. However, Frege’s conception of the generality involved in our discoveries about the objects that are F (or G) is quantificational. Our grasp of generality is exhibited by our characteristic endorsement of inference patterns that abstract from specific, discrete instances: if we discover in the world that *a* is F, then we are entitled to infer that *not every* object is *not* F, or equivalently that *there is* an F.

To judge “there is exactly *one* F” is to judge that there is only a single object, *a*, that is F. To judge “there are exactly two Gs” is to judge there are objects *b* and *c* such that both *b* and *c* are G, and  $b \neq c$ . Together with the claim that no F is a G, we judge that there are exactly these non-identical objects that are F-or-G: *a*, *b*, and *c*. But now, having earlier concluded that there are not three F-or-Gs, we find that our *arithmetical* statement about the *number* of F-or-Gs and

---

temporarily suppose, contrary to what we intuit, that the curved lines on the sphere we see are straight), it serves to make the idea of a non-Euclidean space intelligible.

<sup>72</sup> I owe this point to Tom Ricketts. Frege knew that arithmetic could be axiomatized in second order logic, by using Leibniz’s Law to define equality, using a 1-1 successor operator upon 0 to inductively define the numbers, and inductively defining a multiplication and an addition operation upon the numbers. The numbers are a discrete infinite series. But if we were to leave out one of these requirements—dropping the axiom that 0 is not a successor, for instance, and so allowing finite cyclic models—the resulting structure cannot be manipulated in the ways that characterize our use of arithmetic (for example, it could not be used to model the activity of adding things to a pile, since, at some point, one would not end up with a quantitatively larger number of things). Although we may not contradict ourselves by choosing to define operators in this way, our subsequent reasoning using them does not obviously count as arithmetical.

our *quantificational* statement about the *objects* which are F-or-G have come apart from one another. Denying an arithmetical proposition thus contradicts the equivalence between our arithmetical statements and our quantificational statements.<sup>73</sup>

But, why does Frege think that this sort of contradiction throws *all* of our thought into confusion, rather than merely causing a local problem for our explicit thoughts about numbers of things? I think that the answer to this is tied to Frege's purpose in designing the *Begriffsschrift*, namely, to create a language in which any instance of valid scientific reasoning can be clarified. Frege believes that formalization of a theorem within *Begriffsschrift* will make explicit any hidden premises upon which scientists rely for its proof. But formalizing deductions within *Begriffsschrift* exposes the extent to which scientists implicitly depend upon quantificational inference when reasoning. Since the quantificational generality that undergirds quantificational inference presupposes the existence of discrete objects being generalized over, to meet his goal of designing a language in which scientific reasoning can be clarified, Frege is forced to recognize discrete objects. He accordingly builds discreteness into his conception of "objecthood," and adopts Leibniz's Law as the basic identity criterion for objects (FA, §65). He requires first-level concepts to be well-defined functions for all objects, and licenses second-order quantificational generalizations by appealing to the identity of the value-ranges that go proxy for those concepts under Basic Law V. This process can in principle be repeated higher up the hierarchy of concepts, serving in turn to ground higher-order generalizations. But this entire quantificational hierarchy comes crashing down if one denies an arithmetical proposition,

---

<sup>73</sup> In conversation, Steve Awodey has pointed out that Frege's logicism provides a more straightforward reason for the difference between arithmetic and geometry. Since Frege thinks that each arithmetical truth has a purely logical proof, denying one would immediately result in logical contradiction. But at this point in *Foundations*, Frege has yet to demonstrate logicism. He must therefore be appealing to an independent intuition that he thinks his readers ought to have. So, while Awodey is correct that Frege would have believed the intimate relationship between logic and arithmetic to be ultimately explained by logicism, I contend that here he is thinking of the connection between arithmetical and quantificational statements that he takes to be evident.



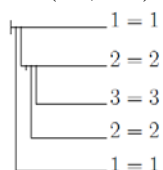
because in losing our ability to *count* objects as distinct numerical identities we lose our grip on their *discreteness*.<sup>74</sup> Since formalizing deductions in the Begriffsschrift exposes that quantificational inferences suffuse scientific inquiry, *all* of our thought is thrown into confusion if we cannot rely upon arithmetic as the stable foundation of the quantificational hierarchy. Completing hypothetical deductions having denied an arithmetical truth is impossible because we have lost the capacity to clearly articulate the thoughts that we wish to grasp, and subsequently judge true.

A more straightforward problem arises in the logical case. When modeled (like the non-Euclidean geometer) within the Begriffsschrift, one who attempts to hypothetically reason having denied a logical law is aiming to make rational inferences on the basis of a series of conditionals, the antecedents of which are the conjunction of the negation of some Begriffsschrift axiom *A* with all of the other axioms, and the consequents of which are substantive theorems. But the Begriffsschrift axioms codify the most general norms of our practice of judgment. To count as judging anything at all we must hold the axioms true. So the deviant who constructs a conditional whose antecedent is the negation of *A* immediately contradicts herself.<sup>75</sup> *Whatever* her chosen consequent, the resulting conditional will be vacuously true. Since logic does not discriminate between the conditionals she may thus construct, her attempt to string together vacuous utterances is itself empty and does not constitute an inference.

---

<sup>74</sup> Since aspects of a single continuous object may appear different to us, we cannot rely upon our ability to distinguish differences (a capacity we have that is presumably independent of our grasp of number) to establish the discreteness of objects.

<sup>75</sup> Frege does admit that a sentence which expresses an instance of the negation of a logical law expresses a genuine thought, since, although it cannot be asserted without immediate contradiction, it is a component part of thoughts which may be legitimately asserted within the notation (CP, 405). An instance of [1] gives a simple example:



The reason that we can engage in further inquiry with the quack who insistently makes wild inferences about physics is that, so long as we avoid the topic of physics and its subsidiaries, she is recognizable as a member of our inferring community. We find her intelligible as a “physics-defective” judge. But logic counts *every* other topic as a subsidiary, meaning that we cannot engage in any further inquiry with a person who resolutely denies a logical law; such a person is unintelligible as a judge.

We can now see that Frege’s constitutive view is the product of holding firstly that logical truths are maximally general, and secondly, that scientific laws are such that they must be recognized by inquirers if those inquirers are to count as engaging in whatever domain of inquiry those laws govern. Since Frege believes that logic is the maximally general science, its laws express the constitutive norms of our practice of inquiry. Any competent judge must recognize them as true.

Where does Frege’s constitutive view fit into Burge’s pragmatic Euclidean reading? Recall that the crux of Burge’s interpretation is maintaining that the Begriffsschrift is a fallible codification of the independently existing, objective laws of logic. Frege’s constitutive view makes engaging in our inferential behavior a necessary condition for being a judge. But to cash out what it means to “engage in our inferential behavior” we must reference a set of norms to which those who count as engaged in our practice are properly sensitive. So Burge faces a choice: are the relevant norms those that Frege captures in the Begriffsschrift or those stemming from the ontologically objective laws of logic?

Under Burge’s account of self-evidence, although Frege *hopes* that his codification expresses the self-evident laws of logic, he accepts that some laws of Begriffsschrift (particularly the worrisome Basic Law V) may be rationally deniable—not only because they may not be

obvious to an inquirer who finds the concepts they invoke unfamiliar, but because they may not even be *true*. It cannot, then, be that endorsing the norms grounded in *this* codification of logic is a constitutively necessary criterion for thought. Indeed, Burge considers it a strength of his interpretation that his Frege would not uncharitably denounce critics of Begriffsschrift as non-thinkers who have failed to grasp that his codification is self-evident. Rather, Burge's Frege would seriously engage his critical interlocutors and, were they proven correct, work with them toward the shared goal of developing a codification of the actually self-evident, ontologically objective laws of logic.

Hence, under Burge's interpretation, Frege's constitutive view is a claim about the norms which are grounded in the objective logical laws. These are the laws which cannot be sincerely denied by a fully rational thinker and which all thinkers must display some grasp of in their inferential practice. In fact, Burge thinks Frege's constitutive view enters his epistemology by providing another sort of justification. In addition to our pragmatic justification for holding the laws expressed in the Begriffsschrift true, we now have a justification "entitling" us to hold the objective laws of logic themselves true:

[Frege holds that] acknowledgment of certain laws of truth is necessary for having reason and for engaging in non-degenerate thinking and judging. One is rationally entitled to judge the primitive laws of logic to be true because the nature of reason—and even non degenerate judgment—is partly constituted by the prescription that one acknowledge at least the simple and basic laws of truth. To put it crudely, reason and judgment—indeed mind—are partly defined in terms of acknowledging the basic laws of truth. (2005, 315-316)

Burge goes on to say that "the laws of truth are independent of judging subjects" and that they are not "constitutively dependent on...practice" (316). This is a consequence of B1. How, then, do we know that we *are* displaying some grasp of those laws in our practice? What justifies thinking that what we *hold* to be laws of truth *are* laws of truth?

Once again, Burge turns to pragmatism to answer both of these questions on Frege's behalf. According to him, Frege thinks that reflecting upon our inferential practice not only

uncovers objective laws of logic, but pragmatically justifies us to hold that the laws we thus uncover *are* objective truths. But the problem with this account is that whatever certification we can attain for our codification of logic by reflecting upon our inferential practice must be fallible, and so our confidence in the truth of our codification necessarily falls short of certainty. In other words, we can never be sure that we have a grasp of the objective laws of logic. This has the effect of forever alienating us from the correctness of our own inferential practice.

To see this, suppose we all agree that a given codification deeply captures the norms of our inferential practice. Suppose we also agree that our practice *seems* good (it seems to allow us to establish truths, it seems to allow us to negotiate the world, and so on). Nevertheless, on Burge's view our judgment that the codification expresses the ontologically objective laws of logic is fallible. Suppose the codification seems undeniable. We take putative objections to call the rationality of the critic into question, not the codification. Although it may seem plausible that we have now captured some of the constitutive laws of thought in the formalism, our imagination is similarly dubitable. We could be badly mistaken.

Since Burge thinks that practice-independent laws of logic bear the constitutive relationship to cognition, the fact that we cannot be certain that we are displaying a grasp of these laws in our practice means that we cannot even be certain that *we* are judges. Our subjective *impressions* that we are indeed judging carry no infallible justification that we are grasping ontologically objective laws. However extreme our confidence in the truth of laws that are particularly basic to our inferential practice—such as the law that every object is self-identical—our confidence is justified only by our fallible judgment in the efficacy of our inferential practice in uncovering truths. To Burge's Frege, objective laws of logic are to what we tentatively judge our inferential practice to be subject.

Frege's unflinching attitude toward the second species of logical alien brings the flaw in Burge's reading into sharp focus. If we were to meet beings who unrepentantly rejected some norm of our inferential practice, while appearing to jabber away at each other quite happily, Frege tells us that we should not be shaken. We are prevented, he says, from supposing that they are right in rejecting our law, and also prevented from doubting whether we or they are right (GGZ1, xxvii). We are entitled, that is, to immediately judge the aliens wrong, and should experience *no* doubt that we might be wrong. On Burge's interpretation, this attitude is inexplicable and unjustified. Upon encountering the second species of alien, the best Burge's Frege can do is to stubbornly insist that *we shall hold their practice to be wrong* because the laws that we are following *seem* right to us. But it nevertheless remains an open possibility that the *aliens* are right about the ontologically objective laws of logic, and that we are very badly mistaken. Since the correctness of our codification of logic is fallibly certified (but not guaranteed) by our pragmatic evaluation that our scientific inquiries are generally successful, we cannot close off this skeptical possibility. We are left foot-stamping: it becomes an undefended assumption that we are rational beings who are equipped with an ability to accurately grasp the objective laws of logic, while the aliens are mad (even though, on this picture, it seems hard to deny that those aliens may have arrived at a parallel verdict about our own sanity). Recall that Burge correctly notes that Frege was singularly unconcerned by skepticism, valorizing his attitude as part of the "advanced spirit of the age" (2005, 340). But the point is that Burge's interpretation makes it thoroughly mysterious why Frege was *not* concerned by skepticism, given that his discussion of logical aliens so strongly brings out the skeptical threat to his view, and, indeed, invites it. Given their apparently functioning deviant practice, how do we know that it is we, and not the aliens, who have grasped the logical laws?

I think that the root of Burge's mistake is the thought that we have a grasp of truth that is independent of our demonstrative inferential practice. This makes it intelligible to question whether the principles guiding our own inferential practice could in fact be false. But on my reading, Frege thinks our sole understanding of truth comes *through* our engagement in our inferential practice, as our name for our goal as judges. Frege writes that "the laws of logic are nothing other than an unfolding of the content of the *word* true" (PW, 3, my emphasis). Our ability to infer demonstratively, displayed by our recognition of the laws of logic, constitutes our grasp of truth. The Fregean logician—who, as Burge recognizes, develops her codification by focusing on the patterns of inference in which we reflectively engage—is concerned to explicitly identify and codify the norms governing our practice, not to quest for practice-independent objective laws of truth of whose grasp she can never be certain. We do well to recall Frege's admonition: "what logic is really concerned with is not contained in the word 'true' at all but in the assertoric force with which a sentence is uttered" (PW, 252). So the Fregean logician's concern is to clarify how some of our utterances come to carry the weight of *assertion*, how our practice of inquiry structures thoughts so that one may *contradict* another, and why some thoughts *imply* others.

On my reading, Frege's constitutive view comes into the foreground. A person's membership in our inferential practice, demonstrated by her recognition that the laws of Begriffsschrift are self-evidently true (since they make explicit norms of the practice of which she is, *ex hypothesi*, a member), constitute her having access to thought, and having the status of "judge." There is no unanswerable doubt about the "correctness" of our inferential practice to an external and inaccessible standard of truth. Rather, the shared standard for the correctness of our inferential judgments is established by that very practice, as we reach a reflective equilibrium

about which judgments and rules we shall endorse. This makes Frege’s unhesitating attitude to the second alien species not only explicable, but justified. Radical skepticism about our knowledge of logic finds no purchase. We exhibit our certain knowledge of the laws of logic—our grasp of the principles of inference—simply by engaging in our inferential practice.

## **5. Sullivan’s Rejoinder**

Peter Sullivan has recently objected to the sort of interpretation that I have been defending in this chapter. His own view interestingly incorporates insights from both Weiner and Burge. Sullivan accepts, with Weiner, that Frege would not have found semantic arguments that depended upon the meanings of marks within a particular symbolism capable of grounding the normative authority of logic for all thought (2004, 740). Such arguments could only establish that *if* one were to employ the concepts expressed within this symbolism, *then* (to the extent that one is a judge) one ought to abide by the logical principles that are expressed in the symbolism in terms of those concepts. They cannot establish Frege’s further constitutive claim that if one is a judge at all, one *must* grasp the concept of truth and abide by logical laws.<sup>76</sup> Nevertheless, Sullivan denies that Frege’s conception of logic makes essential reference to our demonstrative inferential practice. Furthermore, Sullivan reads Frege as offering his reader an “epistemological justification” for holding non-inferentially known truths—e.g. logical laws of the Begriffsschrift—true. In this final section, after exposing why Sullivan’s attribution of an “epistemological justification” to Frege—much like Burge’s “pragmatic” justification—

---

<sup>76</sup> Despite this grasp of Frege’s universalist conception of logic, Sullivan agrees with Burge that Frege offers genuine semantic arguments for the truth of Begriffsschrift laws (720). In Sullivan’s view, Frege’s universalist conception of logic does not *preclude* a metaperspective from which we can grasp semantic arguments about particular codifications of logic, but (unlike other conceptions) does not *require* that we answer the metaperspectival question of whether our codification of logic can be applied to whichever case presently interests us before we may legitimately forge ahead and apply logic to it (726). As I argue below, I think that Sullivan’s suggestion mistakes Frege’s account of objectivity.

mischaracterizes Frege's intent in developing the *Begriffsschrift*, I will argue that Sullivan is wrong to think that making ineliminable reference to our demonstrative inferential practice is in tension with Frege's conception of logic. This critical work is helpful because it enables us to draw out how twentieth-century developments about how to understand our linguistic practice make Frege's unyielding attitude toward logical deviants crude.

Sullivan's attribution of an epistemological justification to Frege is informed by his reading of the following passage:

Now the grounds which justify the recognition of a truth often reside in other truths which have already been recognized. But if there are any truths recognized by us at all, this cannot be the only form that justification takes. There must be judgments whose justification rests on something else, if they stand in need of justification at all.

And this is where epistemology comes in. (PW, 3)

After noting that Frege does not give any indication here of *how* epistemology is to "come in," Sullivan perceptively exposes a number of puzzles about the sort of work Frege might think epistemology can do. Firstly, it cannot be that an epistemological truth is the ground of the purportedly non-inferentially justified truth, for then it would actually be inferentially justified on the basis of epistemology. We would also be left wondering about the justification of the epistemological truth itself. Furthermore, since Frege writes that logical laws neither need nor admit of proof (FA, §3), what rational contribution to the justification of logic can epistemology provide?

One might think that it is just here that Frege problematically turns to "self-evidence." Epistemology comes in by identifying certain truths as "self-evident"; the epistemologically informed inquirer thus recognizes the laws of logic as instances of truths whose truth is apparent once they have been understood. But Sullivan (like Burge) is keen to defend Frege from the charge of psychologism, and so wishes to avoid a naïve attribution about the role of self-evidence to him. The problem Sullivan has with this naïve account as an interpretation of Frege



is that Frege writes so little about (and never defines) self-evidence, which would be surprising if it was to play such a central role in his overall epistemology. “Self-evidence” would seem only to name a problem plaguing Frege’s foundationalism, rather than providing us with the conceptual resources to resolve it. Beyond interpreting Frege, the problem with this account *as* an account is that it is difficult to articulate within a foundationalist epistemology a substantive notion of self-evidence that is distinct from, and does not reduce to, obviousness.<sup>77</sup>

Instead—and again like Burge—Sullivan appeals to a new sort of justification to explain how epistemology “comes in” to justify logic. Epistemology does not give a psychological and essentially causal account of why those judges with human minds judge logical laws true. Nor does it reveal other, distinctively epistemological truths about the nature of knowledge (whatever the psychological framework of the judge) that stand as the rational ground for logical laws. Instead, epistemology provides *us* with a reason to *hold that* logical laws are true.<sup>78</sup> Sullivan’s idea is that if the epistemologically informed inquirer reflects upon her codification of logic and appreciates “the adequacy of that exposition to its guiding norm of truth,” she will be positioned to recognize not only that she is a rational being who has access to the logical source of knowledge—the ability to recognize of analytic *a priori* truths that they are true—but that *this* codification looks to her to be an “exposition [of] the nature of thought laid out” (741). Her theory of knowledge thus gives her reason to hold that her codification of logic is indeed objectively true, a complete laying out of the nature of inferential justification.

---

<sup>77</sup> Two influential critiques of foundationalism are W. V. Quine’s (1951) attack on the analytic/synthetic distinction and Wilfrid Sellars’ dismissal of the myth of the given (1956). Richard Rorty writes that Quine “helped destroy the rationalist form of foundationalism” while Sellars destroyed “the empiricist form of foundationalism” (1997, 5).

<sup>78</sup> Sullivan writes, “[epistemology] comments on the whole chain [of inferential justification]—as it were, from above” (737). His talk of the new justification coming “from above” closely parallels Burge’s account of the pragmatic justification for Begriffsschrift, which was meant in the first instance to apply to the codification as a whole and then derivatively to each individual self-evident axiom.

But we have already seen the problem with attributing this sort of justification to Frege. Our epistemological theories, just like our scientific theories, are fallible. Any judgment we make about the extent to which our codification lays out the nature of thought is one more fallible part of our theory, which cannot therefore make us certain that we *are* grasping objective logical truth. Yet Frege clearly believes that we *do* have a certain grasp of logical laws as self-evidently true. Sullivan's account of how epistemology "comes in" cannot, then, be what Frege had in mind.

In my view, Frege does not think that epistemology "comes in" by giving us a reason or a justification for anything. He makes no such suggestion in the quote upon which Sullivan builds his interpretation. Rather, in Frege's view, the epistemologically informed inquirer is one who grasps that there are truths that all judges must know simply in virtue of being judges. For every judge, qua judge, grasps objective truth as the goal of their demonstrative inferential practice, and manifests their respect for the laws of logic (which, again, merely unfold the content of the word "true" [PW, 3]) in making inferences. Sullivan is correct that Frege gives no substantive account of "self-evidence," and provides no test for discerning whether a particular truth is self-evidently true. But this lacuna is unsurprising given that Frege's aim in developing *Begriffsschrift* is neither to establish the foundations of knowledge nor to reveal and argue for the truth of previously unrecognized truths. Rather, he merely aims to improve and clarify the inferential practice in which he supposes that his audience already (though perhaps imperfectly) engages. Frege does not think that knowing about the nature of knowledge will *justify* our knowledge of logical truths, but only that it will contextualize the constitutive role that logic has to our practice of inquiry.

However, Sullivan also raises an interesting (though confused) objection to any interpretation of Frege's conception of logic, such as my own, which makes ineliminable reference to our demonstrative inferential practice:

Grasp of the kind of norm [that] truth is excludes any...essential reference to "linguistic practice", as much as, and in the same way as, it excludes reference to human psychological processes. (footnote) That Frege endlessly complains against the second while hardly mentioning the first is indicative only of the kind of misunderstanding he encountered amongst his contemporaries. (717)

Sullivan's anxiety is that emphasizing our inferential practice allows contingent claims about us to intrude into logic, resulting with a conception of logic—call it *sociologism*—that cannot account for our grasp of "the kind of norm truth is." Since the purported arguments against sociologism are merely *mutatis mutandis* variations of Frege's tireless arguments against psychologism, Sullivan concludes that the sociological conception of logic cannot be Frege's, despite Frege's "hardly mentioning" the problem of making essential reference to our linguistic practice in discussing truth.

To evaluate this criticism, we must get clearer about two questions. Firstly, does sociologism share whatever feature of psychologism it is that Frege believes prevents the psychological logician from capturing the kind of norm that truth is? Secondly, does the interpretation that I have been defending commit Frege to sociologism?

In section four we saw that Frege argues that since psychological logicians view logical laws as empirically discoverable generalizations about what most human judges hold true, their logical laws only issue in conditional prescriptions for those who wish to judge in accordance with the majority of human judges. Were human brain chemistry to undergo a radical alteration, the psychological law of non-contradiction might have to be replaced by a law of contradiction. In contrast, Frege insists that genuine logical laws issue in unconditional prescriptions for all

judges, because they unfold the content of the word true, which is the constitutive goal of the practice of judgment.

The objectionable sociologism worrying Sullivan, then, would be to view logical laws as empirically discoverable generalizations about what most judges who engage in our linguistic practice hold true. Is the sociological logician forced to accept that her logical laws only issue in conditional prescriptions for those judges who wish to engage in *our* practice of judgment, and that, if our inferential practice undergoes a surprising shift so that people begin freely contradicting themselves, our “sociological” logical laws would alter? So described, this position certainly seems foreign to Frege, and in particular unable to capture his insistence that the logical laws are “eternal boundary stones” for our thought.

To begin defusing this objection, it will help to pick up Sullivan’s remark that Frege “hardly mentions” the evident problems of making essential reference to our practice in developing a conception of logic. This implies that Frege does make *some* mention of these problems—but where? Although Sullivan does not say, a related objection from Burge (2005, 309) suggests that he may be thinking of a passage where Frege criticizes Achelis’ claim that

[discovering] the norms which hold in general for thinking and acting...[requires] an empirico-critical determination of the objective principles of our psycho-physical organization which are valid at all times for the great consciousness of mankind. (quoted in PW, 146-147)

In rehearsing his argument against the psychologism evidenced by Achelis’ view, Frege writes in terms that might seem broad enough to constitute a criticism of Sullivan’s imagined sociologism. In particular, Frege says that were logical norms “norms only because we seldom deviate from them, if it [were] normal to judge in accordance with our laws of logic as it is normal to walk upright,” then since what we “normally” do can change (just as our ancestors did not walk upright), we should have to ask (absurdly, in his view) whether logic can change. It is not difficult to hear this as an objection not only to psychological logicians who hold that we

“seldom deviate” from logical laws because of our brain chemistry, but also to sociological logicians who hold that we “seldom deviate” from logical laws because we have learned to engage in certain patterns of inferential behavior with each other.

Yet I think that this would be to take Frege’s objection out of its context and to mistake its force. Frege’s objection to psychological logicians like Achelis is that they think that logical laws derive their authority from how we actually think and judge. Sullivan assumes that to emphasize the role of our inferential practice in a conception of logic is likewise to insist that logical laws derive their authority from our actual inferential behavior. This sort of sociologism would indeed face a parallel objection to psychologism. Its proponents would owe us an account of why logical laws do not shift with changes in our inferential practice. But on my reading, this is not how our inferential practice enters into Frege’s conception of logic, and so Sullivan’s attack fells a straw man. What Sullivan misses is that, for Frege, the only independent standard of objectivity that a purported codification of logic must meet—the only independent standard of objectivity that there *is*—are the patterns of inference that we reflectively endorse as *exhibited* in our inferential practice. Frege does not think, as the sociological logician does, that logical laws are objectively true (and derive their authority over our thought) *because* they are implicit in our inferential behavior. Rather, he thinks that logical laws are implicit in our inferential behavior (and have authority over our thought) *because* they are objectively true—that is, because they unfold the shared, objective norm that is the constitutive goal of our demonstrative inferential practice. In Frege’s view, if for some reason our behavior systematically changed so that we began to freely contradict ourselves, logical laws would not shift to match our new behavior. Our core logical laws are in reflective equilibrium and would stand firm in their eternal

foundation. Depending on the extent of our inferential malady, it is we who would have become insane.<sup>79</sup>

To put the point another way, Sullivan is correct to insist that objective truth is a primitive notion for Frege's universalist conception of logic (737), but is wrong to think that Frege believes our practice of demonstrative inference is only good *insofar as* it is in accordance with this objective standard. Rather, Frege thinks that we have no grasp upon objective truth outside of or apart from success within our inferential practice of judgment, which Frege insists is not merely *our* practice of judgment—one genus of a larger species—but is rather *the* practice of judgment.<sup>80</sup> We demonstrate our grasp of objective truth by engaging in the practice of judgment, and it acts as a firm standard for all judges.

---

<sup>79</sup> Frege introduces his criticism of Achelis by insisting that we “be wary of the view that it is the business of logic to investigate how we actually think and judge when we are in agreement with the laws of truth,” because in doing so we “should have constantly to have one eye on the one thing and one eye on the other,” and eventually “be seduced into asking questions with no clear meaning” (146). The question which the psychological logician must face is: might all beings with a brain chemistry like ours judge this logical law true, and yet it be false? The question has “no clear meaning” because in seriously asking it the psychological logician reveals that she has a tenuous grasp of truth and does not understand what makes a law “logical.” The question which Sullivan’s sociological logician must face is: might all beings engaged in our inferential practice judge this logical law true, and yet it be false? This question also lacks a clear meaning, and the sociological logician who seriously asks it demonstrates that she has failed to grasp that truth, as the constitutive goal of judgment, is not something judges could mistakenly attribute to basic principles of their inferential practice. In contrast, Frege recognizes that logicians do not need to have their eyes on two things at once—which results in the temptation to ask these questionable questions—but rather *only* have their eyes on our demonstrative inferential behavior, which is the manifestation of our grasp of the objective truth which is the goal of our practice—the practice—of judging.

<sup>80</sup> Sullivan is strangely blind to this point. He recoils from Ricketts’ claim that, to Frege, “the distinction between objective and subjective exhibited in our linguistic practice needs no securing and admits of no deeper explanation” (1986, 72) by questioning the “incongruous” mention of “our linguistic practice” (715). Why, Sullivan wonders, does Ricketts think that the distinction between objective and subjective *needs* no “securing” and yet *gets* it from our inferential practice (as evidenced, according to Sullivan, by Ricketts’ assertion that our linguistic practice “funds” Frege’s conception of logic)? Sullivan grants that we may say that “truth is a norm governing a [linguistic] practice,” but only if we also accept that, since participating in a practice and acknowledging its constitutive norm are one and the same, neither can be understood in terms of the other (717). But this was exactly Ricketts’ point. Since Frege thinks that objective truth *cannot* be understood *except* in terms of our inferential practice, the logician only has our inferential practice to go on in trying to codify the laws of objective truth that inform our practice. Frege’s conception of logic is “funded” by our linguistic practice, not in the only senses that Sullivan, in retaining the fiction of an independent standard for objective truth, can understand of one thing “grounding, supporting, or justifying” (717) a *distinct* thing (i.e., that engaging in our linguistic practice is the *ground* of our grasp of objective truth), but by a helpful redescription of itself in terms of its constitutive goal (i.e. engaging in our linguistic practice *is to* grasp objective truth).

In conclusion, on my interpretation, Frege is not committed to the sociological logician's view of objectivity, under which we think ourselves justified in taking ourselves to have grasped objective truth in our codification of logic *because* that codification seems to capture the current norms that guide our demonstrative inferential practice. Rather, our grasp of objective truth is manifest in our inferential practice, and our aim in developing a codification of logic is to elucidate the norms which govern our practice.

Nevertheless, in drawing out the role that our practice does play in Frege's conception of objectivity, Sullivan's objection serves to put pressure upon Frege's unflinching attitude toward logical aliens, and the immutability of our practice of judging. For the second species of aliens whom Frege dismissed as mad belonged to a community who initially appeared to us to be engaged in a *practice* (GGZ1, xvi). Mightn't their practice—which they seem to meaningfully engage in with each other—exhibit their grasp of a different constitutive goal to our own? Alternatively, might there be a common structure to part of our mutually incompatible practices, meaning that the aliens have some grasp of objectivity and some capacity for cognition after all? What justifies us in reckoning the aliens mad or non-rational, as opposed to concluding merely that we have yet to find a method of communicating with them? Mightn't the aliens' behavior suggest new patterns of inference to us and cause unprecedented shifts in our own practice of judgment? In the wake of the linguistic turn, and the new attention paid to translation of languages and communication between language users, Frege's views seem crude. In the following two chapters, I shall investigate two more recent attempts to capture the relationship between logic and cognition in the work of Quine and Davidson.

## **Bibliography**

- Beaney, Michael, ed. *The Frege Reader*. Oxford: Blackwell, 1997.
- Bridges, Douglas. "Constructive Mathematics." *The Stanford Encyclopedia of Philosophy* (Winter 2009 Edition) Ed. Edward Zalta. URL = <http://plato.stanford.edu/archives/win2009/entries/mathematics-constructive/>.
- Brouwer, L.E.J. "Intuitionism and Formalism." Trans. A. Dresden. *Philosophy of Mathematics: Selected Readings* Eds. Paul Benacerraf and Hilary Putnam, 1983. 77-89.
- Burge, Tyler. "Frege on Knowing the Foundation." *Mind*, 107 (1998): 305-347. Rpt. Burge 2005, 317-355.
- \_\_\_\_\_. "Frege on Knowing the Third Realm." *Mind*, 101 (1992): 633-649. Rpt. Burge 2005, 299-316.
- \_\_\_\_\_. "Frege on Truth." in Haaparanta and Hintikka 1986, 97-154. Rpt. Burge 2005, 83-132.
- \_\_\_\_\_. "Postscript to 'Frege on Truth.'" in Burge 2005, 133-152.
- \_\_\_\_\_. *Truth, Thought, Reason: Essays on Frege*. Oxford: Oxford UP, 2005.
- Conant, James. "The Search for Logically Alien thought: Descartes, Kant, Frege and the Tractatus." *Philosophical Topics*, 21, 1 (1991): 115-80.
- Currie, Gregory. "Was Frege a Linguistic Philosopher?" *British Journal for the Philosophy of Science*, 37 (1986): 79-92.
- Descartes, René. *Discourse on Method and Meditations on First Philosophy*. [1637/1641] Trans. Donald Cress. Fourth ed. Indianapolis: Hackett Publishing Company, 1998.
- Dummett, Michael. *Frege and Other Philosophers*. Oxford: Oxford UP, 1991.
- \_\_\_\_\_. *Frege: Philosophy of Language*. London: Duckworth, 1973.
- \_\_\_\_\_. *The Interpretation of Frege's Philosophy*. London: Duckworth, 1981.
- Evnine, Simon. "Frege on the Relations between Logic and Thought." Unpublished Manuscript.
- Floyd, Juliet and Sanford Shieh, eds. *Future Pasts: The Analytic Tradition in Twentieth Century Philosophy*. Oxford: Oxford UP, 2001.
- Frege, Gottlob. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. (BGS) Trans. S. Bauer-Mengelberg in van Heijenoort, 1967, 1-82.
- \_\_\_\_\_. *Collected Papers on Mathematics, Logic, and Philosophy*. (CP) Ed. Brian McGuinness. Trans. Max Black, V.H. Dudman, Peter Geach, Hans Kaal, E.H.W. Kluge, Brian McGuinness, R.H. Stoothoff. Oxford: Blackwell, 1984.
- \_\_\_\_\_. *Conceptual Notation and Related Articles*. (CN) Trans. and Ed. Terrell Ward Bynum. Oxford: Clarendon, 1972.



- \_\_\_\_\_. *The Foundations of Arithmetic*. (FA) [1884] Trans. J. L. Austin. Second ed. Oxford: Blackwell, 1980.
- \_\_\_\_\_. *Grundgesetze der Arithmetik*. (GGZ1) [1893] Vol. 1. Trans. Michael Beaney in Beaney 1997, 194-223; Trans. Montgomery Furth as *The Basic Laws of Arithmetic: Exposition of the System*. Berkeley: University of California Press, 1964.
- \_\_\_\_\_. *Grundgesetze der Arithmetik*. (GGZ2) [1903] Vol. 2. Trans. Peter Geach and Michael Beaney in Beaney 1997, 258-289.
- \_\_\_\_\_. *Philosophical and Mathematical Correspondence*. (COR) Eds. Gottfried Gabriel, Hans Hermes, Friedrich Kambartel, Christian Thiel, and Albert Veraart. Trans. Hans Kaal. Chicago: University of Chicago Press, 1980.
- \_\_\_\_\_. *Posthumous Writings*. (PW) Eds. Hans Hermes, Friedrich Kambartel, and Friedrich Kaulbach. Trans. Peter Long and Roger White. Chicago: University of Chicago Press, 1979.
- \_\_\_\_\_. "Thought" (T) [1918-1919] Trans. Peter Geach and R. H. Stoothoff in Beaney 1997, 325-345.
- Goldfarb, Warren. "Logicism and Logical Truth." *Journal of Philosophy*, 79 (1982): 692-695.
- \_\_\_\_\_. "Frege's Conception of Logic." Floyd and Shieh 2001, 25-42. Rpt. Potter and Ricketts 2010, 63-85.
- Greimann, Dirk. "Does Frege Use a Truth-Predicate in his "Justification" of the Laws of Logic? A Comment on Weiner." *Mind*, 117 (2008): 403-425.
- Haaparanta, Leila and Jaakko Hintikka, eds. *Frege Synthesized*. Dordrecht: D. Reidel, 1986.
- Heck, Richard. "Frege and Semantics." Potter and Ricketts 2010, 342-378.
- Jeshion, Robin. "Frege: Evidence for Self-Evidence." *Mind*, 113 (2004): 131-138.
- \_\_\_\_\_. "Frege's Notions of Self-Evidence." *Mind*, 110 (2001): 937-76.
- Kolodny, Niko and John MacFarlane. "Ifs and Oughts." *Journal of Philosophy*, 107 (2010): 115-143.
- Kremer, Michael. "Judgment and Truth in Frege." *Journal of the History of Philosophy*, 38 (2000): 549-581.
- Linnebo, Øystein. "Frege's Conception of Logic: From Kant to Grundgesetze." *Manuscripta*, 26, 2 (2003): 235-52.
- MacFarlane, John. "Frege, Kant, and the Logic in Logicism." *The Philosophical Review*, 111 (2002): 25-65.
- Massey, Barbara. "Do All Rational Folk Reason As We Do? Frege's Thought Experiment Reconsidered." *Thought Experiments in Science and Philosophy*. Eds. Gerald Massey and Tamara Horowitz. Savage, MD: Rowman and Littlefield, 1991. 99-112.

- Potter, Michael and Thomas Ricketts, eds. *The Cambridge Companion to Frege*. Cambridge: Cambridge UP, 2010.
- Proops, Ian. "Russell and the Universalist Conception of Logic." *Noûs*, 41 (2007): 1-32.
- Quine, W.V. "Two Dogmas of Empiricism." *The Philosophical Review*, 60 (1951): 20-43.
- Ricketts, Thomas. "Concepts, objects and the Context Principle." Potter and Ricketts 2010, 149-219.
- \_\_\_\_\_. "Frege, the Tractatus, and the Logocentric Predicament." *Noûs*, 19 (1985): 3-15.
- \_\_\_\_\_. "Frege's 1906 Foray into Metalogic." *Philosophical Topics*, 25 (1997a): 169-88.
- \_\_\_\_\_. "Logic and Truth in Frege." *Proceedings of the Aristotelian Society*, 70 (1996): 121-40.
- \_\_\_\_\_. "Objectivity and Objecthood: Frege's Metaphysics of Judgment." Haaparanta and Hintikka 1986, 65-95.
- \_\_\_\_\_. "Truth-Values and Courses-of-Value in Frege's *Grundgesetze*." in *Early Analytic Philosophy: Frege, Russell, Wittgenstein*. Ed. W.W. Tait. Chicago: Open Court Press, 1997b. 187-211.
- Reck, Erich, ed. *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy*. Oxford: Oxford UP, 2002.
- Rorty, Richard. "Introduction." *Empiricism and the Philosophy of Mind*, Wilfrid Sellars. Cambridge: Harvard UP. 1997, 1-12.
- Sellars, Wilfrid. "Empiricism and the Philosophy of Mind." *Minnesota Studies in the Philosophy of Science*, vol. 1. Ed. Herbert Feigl and Michael Scriven. Minneapolis: University of Minnesota Press, 1956.
- Shieh, Sanford. "On Interpreting Frege on Truth and Logic." Reck 2002, 96-124.
- Sluga, Hans. "Frege: The Early Years." *Philosophy in History*. Ed. Richard Rorty, J.B. Schneewind and Quentin Skinner. Cambridge: Cambridge UP, 1984. 329-356.
- Sullivan, Peter. "Frege's Logic." *Handbook of the History of Logic*. Eds. Dov Gabbay and John Woods. Vol. 3. Amsterdam: Elsevier/North Holland, 2004, 671-762.
- \_\_\_\_\_. "Metaperspectives and internalism in Frege." *Gottlob Frege: Critical Assessments of Leading Philosophers*. Vol. 2. Eds. Michael Beaney and Erich Reck, London: Routledge, 2005. 85-105.
- Tappenden, Jamie. "Metatheory and Mathematical Practice in Frege." *Philosophical Topics*, 25 (1997): 213-64.

- Taschek, William. "Truth, Assertion, and the Horizontal: Frege on 'The Essence of Logic'." *Mind*, 117 (2008): 375-401.
- Textor, Mark. "A Repair of Frege's Theory of Thoughts." *Synthese*, 167 (2009): 105-123.
- van Heijenoort, Jean. "Logic as Calculus and Logic as Language." *Synthese*, 17 (1967): 324-330.  
 \_\_\_\_\_, ed. *From Frege to Godel: A Sourcebook in Mathematical Logic 1879-1931*.  
 Cambridge: Harvard UP, 1967.
- Weiner, Joan. "Burge's Literal Interpretation of Frege." *Mind*, 104 (1995a): 585-597.  
 \_\_\_\_\_. *Frege in Perspective*. Ithaca: Cornell UP, 1990.  
 \_\_\_\_\_. "How Tarskian is Frege?" *Mind*, 117 (2008): 427-450.  
 \_\_\_\_\_. "Realism *bei* Frege: Reply to Burge." *Synthese*, 102, 3 (1995b): 363-382.  
 \_\_\_\_\_. "Semantic Descent." *Mind*, 114 (2005): 421-354.  
 \_\_\_\_\_. "What was Frege trying to prove? A Response to Jeshion." *Mind*, 112 (2004):  
 115-29.
- Wilson, Mark. "Frege: The Royal Road to Geometry." in *Frege's Philosophy of Mathematics*,  
 Ed. William Demopoulos. Cambridge: Harvard UP, 1995. 108-149.
- Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*. [1921] Trans. C. K. Ogden. London:  
 Routledge, 2002.

## **Quine's Naturalist Conception of Logic: Excising Skepticism from Epistemology**

Language is a social art. In acquiring it we have to depend entirely  
on intersubjectively available cues as to what to say and when.

These opening sentences of W.V. Quine's *Word and Object* signal its central concerns.

Quine aims to understand the nature of language—its acquisition and its use. The conclusions he draws are not limited to semantics. They extend to metaphysics and epistemology. In fact, because of Quine's influential, sometimes maligned, and often misunderstood brand of naturalism, "epistemology merges with psychology, as well as with linguistics" (EN, 90). In this chapter, in dialogue with Quine's critics, my aim is to argue for naturalism's attraction and power by scrutinizing Quine's conception of logic and, in particular, his account of the relationship between logic and cognition.

Naturalism is not a thesis within a branch of philosophy, but rather a conception of proper philosophical activity. In one formulation, Quine writes that it "means banishing the dream of a first philosophy and pursuing philosophy rather as part of one's system of the world, continuous with the rest of science" (RHP, 430). But to espouse naturalism is not to adopt a crass scientism.<sup>1</sup> Instead, naturalism acts as a methodological paring knife. It cuts away a swathe of problems by exposing that their presuppositions are confused and carves out territory in which subsequent philosophical activity can be gainfully pursued.

To some readers, the aforementioned description might suggest that a machete is a more apropos metaphor for naturalism, and naturalists are sometimes cast by their opponents as decimating the philosophical landscape.<sup>2</sup> Nevertheless, I think that, rightly understood,

---

<sup>1</sup> Charges of scientism have been leveled against Quine by a number of philosophers who detect only prejudice in the view. See e.g. Hilary Putnam (1998, 425), and Thomas Sorrell (1991).

<sup>2</sup> Most recently, see Joseph Margolis (2003, 5-8).

naturalism is a nuanced position. The labyrinthine exploration of it that I undertake will allow us to grasp its subtlety, enabling us to critically evaluate it.

In the next section, I extract Quine's view of the relationship between logic and cognition from his account of radical translation. I argue that although he adopts Frege's insight that the logician's task is to codify the principles of *our* demonstrative inferential practice, his conception of logic bears no trace of Frege's constitutive view. This absence raises questions about how Quine thinks inquirers are able to identify each other, and whether he can account for the normative character of our practice of inquiry. To address them, we must elaborate the naturalism that Quine tends to encapsulate in slogans. My method for doing so will be to show how appealing to naturalism allows the Quinean to excise a variety of skeptical challenges. Hence, in section two, I argue against Barry Stroud that naturalists can avoid skepticism about the external world. In section three, by contrasting Quine's philosophy of language with Rudolf Carnap's, I argue that Quine defuses a skeptical challenge about the possibility of communication. In section four, I defend Quine from Gary Ebbs' charge that he cannot account for our dispute-resolving practices and show that he also eliminates a skeptical threat about how our words ought to be used. Having exhibited the naturalist's ability to remove these skeptical attacks from philosophy, I return to the questions that I raise in the first section and reconstruct Quine's account of the objectivity of our inquiry.

### **1. Naturalism, Logic, and Cognition: The Denial of a Pre-Logical Mentality**

Pre-logicity is a trait injected by bad translators. (CLT, 387)

Quine develops his theory of radical translation in the second chapter of *Word and Object*. He imagines a field linguist trying to translate the utterances of a group of natives. The situation is "radical" because the native language is unlike any with which the linguist is

familiar. Deprived of assumptions about its likely grammar (through comparison with other languages in its family), she is forced to rely solely upon her observations. In particular, she must attend to what sounds the natives make and in what contexts they make them.<sup>3</sup> Her project will rely upon minimal (though falsifiable) assumptions about the natives' sensory capacities. She begins to speculate which of the "forces" she sees "impinging on the native's surfaces" cause the natives to make particular noises (WO, 28).<sup>4</sup> When similar circumstances arise in the future, she can test whether she is correct by emulating native sounds she thinks appropriate and evaluating the natives' reactions. If the natives seem to endorse her utterance, she has some confirming evidence for her hypothesis; if they react with confusion, she has some disconfirming evidence.<sup>5</sup> The provisionally confirmed hypotheses collected in this fashion mark the first tentative steps of translation, for now she can correlate the utterances she would make in those circumstances to native noises.

Some surprises in the natives' behavior—including their speech behavior—are tolerable (and are even to be expected). The linguist may, for example, discover that the natives delight in feasting on the meat of an animal which she holds sacred, calling it "delicious" according to her

---

<sup>3</sup> For simplicity, I focus here upon the verbal. If the linguist suspects that the natives also communicate using non-verbal behavior—as we do when, for example, we raise a finger to our lips to indicate the desire for our interlocutor to be silent—then she may add a sign language appendix to her translation manual. Similar extensions are possible if the translator suspects still more esoteric methods of communication. If we switch the case from natives to aliens, an array of unusual methods based upon alien sensory apparatuses becomes possible (for example, an olfactory system that relies upon systematic pheromone emission). If our own senses, or scientific instruments, can reliably detect how the aliens are stimulated, then we can create a workable translation manual even in this more extreme case. See Neil Tennant (1999) for an extended discussion of radical alien translation.

<sup>4</sup> Quine insists that although the linguist may take the native to be talking *about* things in the world, she ought to focus upon the native's sensory stimulations as the relevant locus for her investigation (WO, 31). His reason is that the translator is in a position to assess, on the basis of her behavioral observations, the natives' sensory stimulations "by making allowances for the speaker's orientation and the relative disposition of objects." In contrast, native behavior does not reveal the natives' private (and possibly quite different) ontology. In chapter four, I will discuss Quine's later modification of this position under pressure from Davidson's arguments.

<sup>5</sup> The linguist's translations of endorsement and reprobation in the native tongue are themselves working hypotheses. Quine suggests that if the linguist regularly hears two words in response to her attempts to use the native language then she has good reason to translate these as "Yes" and "No," determining which is which on the basis of her own attempts at using these words in response to some of the native's utterances (WO, 29-30).

manual. The fact that she would never contemplate eating the animal, much less approving of the taste of its flesh, does not call into question her translation of the native word as “delicious.” Rather, she and the natives have a religious disagreement. Although this may well cause problems for future interactions between the linguist and the natives, it does not constitute a *translation* problem.

However, some surprises in native speech behavior are intolerable. Among these, according to Quine, are disagreements about simple logical truths. If the linguist discovers that translating her interlocutors according to her current manual would have them affirming contradictions, Quine tells us that she should throw it out, rather than be tempted to ascribe a “prelogical mentality” to them (WO, 58).<sup>6</sup> Disagreeing about logic is thus unlike disagreeing about religion. It constitutes a translation problem and requires us to search for a new manual.

This attitude recalls Frege’s rejection of logical aliens. Frege’s ground for dismissal is that beings who affirm contradictions fail to meet the constitutive standard for the attribution of cognition: recognition of the logical laws as immutable, eternal “boundary stones” for thought. Frege might count Quine’s advice to seek a new translation manual a friendly amendment to his view. Suppose we are tempted to judge that the aliens whom we have met are like the second species that I discussed in chapter one. Despite their apparent ability to negotiate the world and their seeming to engage in some sort of linguistic practice, their resolve in uttering strings of contradictions to each other is in danger of undermining our attribution of cognition to them. At first, we thought them mad, but now we wonder whether they are, after all, non-rational brutes.

---

<sup>6</sup> This phrase is due to the pioneering anthropologist Lucien Lévy-Bruhl, who, witnessing native behavior he found inexplicable, posited that the natives had a “prelogical mentality” (1924).

Before we write them off, Frege may concede, we should question whether they are *really* asserting contradictions or if the error lies with our attempt to translate them.<sup>7</sup>

But in contrast to Frege, Quine's objection to the linguist ascribing a prelogical mentality is not that in so doing she represents a putative thinker affirming things that no genuine thinker could intelligibly affirm, but simply that she represents the native as affirming something she finds obviously false. Logic has no monopoly on obviousness: just as it is not credible to translate people consistently calling red objects "blue," or as denying that it is raining during a torrential downpour, so it is not credible to translate people as affirming simple contradictions.<sup>8</sup>

So, unlike Frege, Quine does not think logic has a privileged, constitutive relationship to cognition. In Quine's view, our justified beliefs are not grounded upon an immutable logical foundation, but rather form an interconnected web: "our statements about the physical world face the tribunal of sense experience not individually but as a corporate body" (TDE, 41). Confronted with recalcitrant data, it is easy to give up beliefs toward our web's periphery. It is more difficult, and so rarer, to reject the obvious beliefs near the center (both logical and physical). Nevertheless, as we continue to revise our theories for negotiating and predicting the world, any of our beliefs—even those we currently think obvious—may come up for judgment.

Lying underneath the apparent similarity of Frege's attitude to logical aliens and Quine's to prelogical mindedness are two distinct conceptions of logic. Frege's constitutive view casts the logician as searching for the prerequisites of rationality. Deprived of this explanation of

---

<sup>7</sup> Our judgment will depend upon how far their linguistic practice aids them in negotiating the world. If they jabber at length to each other while seeming to have built great cities, designed works of art, developed electronic equipment, and so on, we may judge that we need a new translation manual. If they only occasionally squawk at each other, and engage only in the primitive behavior that we recognize in non-human animals (such as seeking food when hungry), we may judge them non-rational.

<sup>8</sup> Nor are only logical statements privileged by being "stimulus-analytic" for us, eliciting our assent regardless of how we are currently being stimulated (WO, 55). Empirical statements, such as "there are black cats," also have this property.



logic's importance, what role does Quine think logic has vis-à-vis rational inquiry, and how does he characterize the logician's task?

To resolve a dispute about a contested hypothesis, scientists rely upon their shared understanding of its testable consequences. By running experiments, they discover which consequences obtain, and, accordingly, endorse or reject the hypothesis in question. In Quine's view, the naturalist logician aims to codify the principles of inference that underwrite this procedure. Attending to our beliefs about which claims imply which others allows her to offer a formal explication of the relation of "implication" to which disputants are making tacit appeal. This refines our grasp of scientific practice (WO, 226; PL, 100). The logician ought to strive for her codification of logic to be *canonical*. If she succeeds, that is, it will be possible to regiment *any* current theory using her notation (a fact that simplifies the structure and clarifies the concepts of the conceptual scheme of science [WO, 161]). Just as we sometimes paraphrase ambiguous sentences to clarify what we intended to communicate in cases of confusion, so, too, we may paraphrase our theories in the artificial notation to clarify their inferential commitments in cases of dispute.<sup>9</sup> Moreover, since regimenting our theories exposes their ontological commitments, the canonical notation can assist us in "limning...the most general traits of reality," according to our current theories (161).<sup>10</sup>

---

<sup>9</sup> According to Quine, these paraphrasing techniques differ only in degree (WO, 159). The logician's artificial notation is phrased within ordinary language and is not distinct from it. In section three, we will see Carnap's contrasting view that formal calculi are distinct from natural language (and each other).

<sup>10</sup> In Quine's view, ontological questions belong to science, not metaphysics: "what distinguishes the ontological philosopher's concern [from the concerns of scientists] is only breadth of categories...[I]t is scrutiny of...uncritical acceptance of the realm of physical objects itself, or of classes, etc., that devolves upon ontology" (WO, 275). Guided by logic, the ontological philosopher's maxim is that, "to be is to be the value of a variable" (OWTI, 15). For a good discussion of Quine's views here, see Joseph Ullian (2004, 270-272).

The canonicity of logic goes some way toward accounting for its importance for inquiry. Inquirers aim to develop theories, and canonicity makes regimentability a condition of theoryhood. We will not count something that cannot be regimented into our notation a theory.<sup>11</sup>

Moreover, the fact that our canonical logic explicates our implication relation explains its importance when translating other inquirers. Quine believes that the translation of one's interlocutor's logic—like the rest of one's interlocutor's language—is indeterminate (R, 314; 319). In the case of radical translation, there is no manual-independent fact of the matter about the logical form of the native language. Nevertheless, Quine says that one of the linguist's early goals is to ascertain onto which of the native noises she can map our truth-functional operators. One reason is that these operators allow her to recursively parse native utterances. Without them, she would search in vain for a manual capable of translating any of the infinite well-formed expressions of the native language with expressions of her own.<sup>12</sup> A second reason is that we cannot so much as *begin* translating until we have provisionally taken some native noises to be assent and dissent, and a key part of our evidence for doing so will be the natives' reactions to the noises we translate as the truth-functional operators (e.g. the natives should “assent” to all utterances formed by attaching the “negation” noise to utterances from which they “dissent”).

---

<sup>11</sup> In *Philosophy of Logic*, Quine describes the canonical logic to which he thinks all current rational inquiry is accountable: the first order predicate calculus (including identity). This logic is weaker than the systems preferred by his immediate logical predecessors. Unlike Frege's Begriffsschrift, set theoretic membership is excluded. Unlike Bertrand Russell's type-theoretic system, second- and higher-order logical quantification are excluded. But Quine takes his canonical logic to have a decisive advantage over these other candidates in admitting complete proof procedures for validity and inconsistency (PL, 90-91). George Boolos questions why *this* advantage should count as decisive, as opposed to, for example, decidability. His paper is reprinted and extensively discussed in Shapiro (1996). Intuitionistic logic is dismissed by Quine as lacking “the familiarity, the convenience, the simplicity and the beauty of our logic” (PL, 87).

<sup>12</sup> No finite list of paired utterances can suffice as a manual of translation between languages with infinite well-formed expressions. But, one might object, what if the native language has only finitely many well-formed expressions? Although I cannot address this worry in any detail here, I do not think we can count a finite system of sounds a language unless we also admit that animals use language. A Quinean might be led to admit that animals *do* use language for certain investigations, but this admission obfuscates our current goal of exploring the relationship between logic and cognition. In the next chapter, I will argue that Davidson's conception of this relationship *forces* him to conclude that anything worth calling a language must have infinite expressions.

For these reasons, Quine describes radical translation with this much of the native tongue fixed (though subject, like the rest of the linguist's efforts, to future disconfirmation).

In contrast, translating quantifiers requires isolating predicates with quantifiable positions—and fixing which entities can occupy them—in the native language. Quine argues that this can be accomplished in a number of incompatible ways that will nevertheless be consistent with all the data of native speech dispositions (WO, 60-61). Yet, we have an excellent *pragmatic* reason to translate the natives as using our canonical logic (and as predicating properties of entities in our ontology). Since logic makes explicit the implication relation to which we are committed, this translation provides us with a perspicuous framework within which to overcome disputes with the natives, and so, generally promotes smooth dialogue between us. (This pragmatic reason for attributing our canonical logic also holds in our everyday translations of each other).

So pragmatic considerations guide the linguist and the naturalist logician. The naturalist aspires for her codification of logic to be useful to ongoing empirical inquiry, not to lay out the constitutive prerequisites of thinking. Quine writes that it is “a possibility in principle” that “another culture, another species, [could] take a radically different line of scientific development, guided by norms that differ sharply from ours...[and] predict as successfully and thrive as well as we” (R2, 181). These norms include the prescriptions that issue from our logical laws. We cannot know *a priori* that other inquirers' theories have quantificational structure, nor can we infer that other inquirers' canonical logic is the first order predicate calculus with identity. But we have good reason to adopt a canonical logic to clarify our inquiry and to translate others as adopting it, too.

In fact, Quine thinks that *our own* line of scientific development might result in an alteration of the laws of logic. I think that critics have sometimes masked this striking feature of Quine's conception of logic by conflating it with his skepticism about modal claims.<sup>13</sup> Quine's point is not merely that the claim "it is impossible for us to reject a logical law" relies upon a scientifically indefensible notion of impossibility. Nor is it that future logicians might develop alternative codifications of logic that are superior to our own on the grounds of simplicity or clarity (a point with which Frege would agree). Rather, Quine thinks that our future scientific theories may *require* us to change our logical laws.

To see this, consider that Quine's sole doubt about his canonical notation stemmed from quantum mechanics (PL, 85-86; 100). Heisenberg's Uncertainty Principle entails that magnitudes of quantum particles cannot be jointly ascertained. Measuring the location of a quantum particle makes its velocity indeterminate, while measuring its velocity makes its location indeterminate. This indeterminacy is part of our best current physical theory of the world, not an epistemic failing in us. When certain laws of classical logic are applied to propositions involved in reasoning about quantum mechanics, claims become admissible that according to our own theory, paradoxically, cannot be established.<sup>14</sup> To avoid the proliferation of unanswerable questions based upon indeterminable claims, some logicians developed new, non-classical calculi to model quantum reasoning.<sup>15</sup>

---

<sup>13</sup> Stroud, for example, writes of Quine in one breath that "[t]here is no metaphysically guaranteed eternal unreachability or impossibility, even though many of our present beliefs, *including the truth-functional tautologies*, might never in fact be given up during the rest of the history of human life" (1975, 95, my emphasis).

<sup>14</sup> One example is the distribution of conjunctions over disjunctions. The schema " $[p \cdot (q \vee r)] \equiv [(p \cdot q) \vee (p \cdot r)]$ " is a theorem of classical logic. But if "p" is "the particle is moving to the left," "q" is "the particle is in the interval [0,1]," and "r" is "-q," then " $p \cdot (q \vee r)$ " is equivalent to " $p \cdot (q \vee \neg q)$ " and so to "p," while " $(p \cdot q) \vee (p \cdot r)$ " is unknowable since Heisenberg's Uncertainty Principle means that neither p and q nor p and r can be jointly ascertained. See Jan Hilgevoord and Jos Uffink (2011).

<sup>15</sup> See Garrett Birkhoff and John von Neumann (1936) and Putnam (1974).

Although Quine came to dismiss such “quantum logics” because of their lack of familiarity and conservativeness, the case illustrates the sort of considerations that might compel naturalist logicians. Suppose scientists find that they are *unable* to model their reasoning about a subject in terms of their canonical logic. Then, if logicians devise an alternative codification of the principles of inference which explicates these new patterns of reasoning, and into which the scientists can regiment their new theories (in addition to theories capable of explaining all of the phenomena that theories regimented in the old canonical logic could explain), the alternative codification should be canonized.<sup>16</sup> This rational rejection of one or more of our logical laws would be the result of the development of successful scientific theories (so judged by their predictive power) expressible only in a logical system inconsistent with our own. In this way, logic and scientific theories develop in tandem. Regimentability into canonical logic is a condition for theoryhood; but a widely endorsed, predictive, yet currently unregimentable explanatory account of some phenomenon may result in a new canonization.

The claim that our logical and scientific theories develop in tandem means that Quine, unlike Frege, is willing to accept that thinking, talking, reasoning aliens might exist who resolutely deny our logic. They may successfully create theories about the world that we are unable to regiment using our canonical notation, and their language may be currently untranslatable to us because we cannot project our logic onto their speech. Doubtless, we would be baffled by their alien mindedness. Nevertheless, if over time we came to understand their “radically different line of scientific development” (R2, 181), and the implication relation of their canonical logic (however different from our own), we might be able to reject our logic in

---

<sup>16</sup> Hartry Field (2008) argues that a multi-valued logic, weaker than classical logic, averts the semantic paradoxes. All current scientific theories can be regimented in terms of this logic. Were a Quinean to take Field’s treatment of the paradoxes as preferable to Quine’s own disavowal of them as untranslatable foreign language (PL, 45), we should have another case of considerations that would lead her to revise her canonical logic.

favor of theirs and translate them in the future. Today's logical aliens may be tomorrow's scientific collaborators.

There is a famous passage that might be thought to discredit the interpretation of Quine I have been pursuing:

[Consider] the familiar remark that even the most audacious system-builder is bound by the law of contradiction. How is he really bound? If he were to accept contradiction, he would so re-adjust his logical laws as to insure distinctions of some sort; for the classical laws yield all sentences as consequences of any contradiction. But then we would proceed to reconstrue his heroically novel logic as a non-contradictory logic, perhaps even as familiar logic, in perverse notation. (WO, 59)

Numerous critics have taken Quine's point to be that, try as one might, one cannot help but be bound by the classical logical constants if one is to construct a theory (e.g. Alan Berger, 1990). On this reading, Quine is developing a powerful argument that the law of contradiction is rationally incontrovertible, since one cannot deny it and still construct predictive theories. But what Quine actually says is that we—i.e. classical logicians—*would* reconstrue the alien logic in our own terms. By this, Quine means that when devising our manual of translation for non-classical logicians, we would use our canonical logic to describe their deviant theories so that we can begin to communicate. Given that our canonical logic explicates the implication relation guiding our practice of inquiry, beginning translation by locating native language onto which our logical structure can be read is an excellent rule for translation. Yet, given the indeterminacy of translation, to begin translating in this way is still a pragmatic convention of *ours*, and does not reveal an “essential” truth about the role of classical logic in the alien tongue, or the centrality of classical logic to all languages, or all theories. Rather, it only shows that classical logic is *currently* central in *our* web of belief and currently indispensable for our translation projects. Quine writes that in radical translation,

‘Save logical truth’...is a rule which, compatibly with all stimulus meanings and other verbal dispositions, *could be obeyed or flouted*. But it is not capricious. The very want of determinacy puts a premium on adhering to this strong and simple rule as partial determinant. (R, 318, my emphasis)

We can now state Quine's *naturalist* conception of the relationship between logic and cognition:

**NAT:        A person must exhibit speech behavior that can be translated as reasoning in conformity with the laws of our canonical logic if we are to engage in inquiry together.**

Like Frege, Quine takes the logician's goal to be arriving at an explicit account of the principles of demonstrative inference that regulate rational inquiry. He also agrees with Frege that the principles of inference are *our* principles of inference; the logician's codification is not accountable to a standard that is independent of our demonstrative practice. But unlike Frege, Quine lacks a constitutive view of the relationship between logic and cognition. Although he agrees with Frege that the logician's codification must certify those arguments informally recognized as valid and no others, Quine takes seriously the possibility that what we informally recognize as valid may change with ongoing scientific development. I argued in chapter one that Frege thought his basic logical laws express general truths of rational inquiry that are in reflective equilibrium; but in Quine's view, once the naturalist logician accepts that her codification is just a snapshot of the principles of inference currently accepted in her community, she has no basis for judging any part of her codification eternally fixed.

This naturalist account of the relationship between logic and cognition faces a number of questions. If there are no constitutive norms that *must* be recognized if one is to be a rational inquirer, then how does one know which norms other inquirers recognize? If our norms differ from our interlocutor, are we able to agree and disagree or do we merely talk past each other? In fact, what conception does the naturalist even have of "norms" and "normativity"? Is it sufficient to model our practice of inquiry? And, if we accept a broadly Fregean account of the intimate relationship between logic and truth, then, having accepted that even the core laws of

our canonical logic can be rationally revised, must naturalists accept a form of sociology where objective truth changes with the patterns of reasoning our canonical logic enshrines?

Many of Quine's critics have charged that he lacks adequate responses to some or all of these questions. But such critics fail to appreciate the responses Quine's naturalist conception of philosophy makes available. To show how Quine can address his critics, in the next three sections I want to examine naturalism more carefully. My method will be to expose how appealing to naturalism allows Quine to quell different skeptical objections. Once we have acknowledged naturalism's strengths, we will be positioned to critically evaluate its triumphs and failures.

## **2. Barry Stroud and Skepticism about the External World**

We can develop our understanding of naturalism by considering how, and why, Quine thinks that appealing to it can defuse the problem of skepticism about the external world. Traditional skeptical arguments purport to show that we lack knowledge—or, alternatively, justified beliefs—about the world. But in Quine's view, traditional "*theory of knowledge...blushes for its name*" (RA, 321, my emphasis), since much of its terminology fails to pass scientific muster: "for scientific or philosophical purposes the best we can do is give up the notion of knowledge as a bad job and make do rather with its separate ingredients" (Q, 109). My goal in this section is to argue that, by making do with naturalized epistemic ingredients, Quine removes the presuppositions of traditional skeptical arguments. To do so, I will engage Stroud's criticisms of Quine and argue that he mistakes the power of naturalism. This will allow me to begin fleshing out the significance of Quine's slogans that being a naturalist means rejecting "first philosophy" (RHP, 430) and accepting that there is no "cosmic exile" (WO, 275).



Further, it will let me explore the ramifications of doing naturalized epistemology with Quine upon the deck of Neurath's metaphorical boat (SO, 16).

Let us begin with a working account of what it is to naturalize epistemology. Quine characterizes naturalists as accepting that philosophy is continuous with natural science (RHP, 430). In consequence, naturalists believe that there are no supra-scientific methods available to philosophical inquirers that would license a distinctively "philosophical" critique of scientific methodology or of scientific theories of reality. Instead, our ongoing scientific practice is a *self*-correcting, comprehensive endeavor, committed to employing the best ways that we have managed to discover to learn about the world, so judged "best" by the unsurpassed predictive and explanatory power of the theories scientific inquirers have been able to develop. When naturalists turn to epistemology, they acknowledge that empirical psychologists and linguists share their domain of investigation. Each set of inquirers strive to better understand the human mind's capacity for believing, thinking, knowing, and expressing by carefully attending to the available evidence and constructing theories which can account for observed phenomena. Their investigations differ only in degree, not in kind.<sup>17</sup>

Like their traditional counterparts, Quine tells us that naturalized epistemologists seek the "rationale of reification" (TI, 3), an explanation of how we manage to gain knowledge of things in the world on the basis of our sense experience. But he characterizes them as "liberated" by the recognition that they rely on their background understanding of the world to construct their more sophisticated theories of it (RR, 3). This background understanding is no less a *theory* than its sophisticated progeny. Naturalized epistemologists thus view the ordinary things that we commonly take to populate the world—chairs, trees, apples, etc.—on a continuum with the

---

<sup>17</sup> Quine rebukes non-naturalized epistemologists for engaging in theoretical "make-believe" by *imposing* conceptual structure onto epistemology, rather than *discovering* structure by critically attending to the results of ongoing empirical experimentation (EN, 75).

entities that we have posited, comparatively recently, to account for the data garnered by our scientific investigations (electrons, dark matter, genes, and so on). We may better recall the experimental history leading to these exotic posits, but ordinary things are, likewise, posits of ours whose origins are “shrouded in prehistory” (WO, 22).

Moreover, *we inquirers* number among these ordinary posits. When we turn our gaze upon ourselves, our theories tell us that “we” are physical organisms. Beginning “in mediis rebus,” in the midst of her background theory, the naturalized epistemologist’s task is to reconcile this theoretical account of what we are with our conscious experience of ourselves and of the knowledge we have of the world (RPR, 460).<sup>18</sup> The (currently accepted) scientific view is that organisms like us obtain all the evidence relevant to our theories of reality via our senses, and so, naturalized epistemologists hold that we are warranted in believing theories which are able to explain and predict our sensory experience.

The “currently accepted” qualifier is important. It indicates that Quine’s commitment to physicalism—and even to *empiricism*—is subordinate to naturalism. *Current* scientists are empiricists (holding that we only access the world using our sense organs) and physicalists (holding that the world to be explained by our theories is entirely physical). But since the results of future scientific inquiry may necessitate non-physical theories (Quine remarks that some accounts of quantum physics are virtually mentalistic), and since we could evolve a (verifiable) capacity to access the world which did not rely on our senses, orthodox physicalism or empiricism could be abandoned (PT, 20-21). Quine’s commitment to naturalism, though, would remain. Naturalistic philosophy would seek to evaluate the commitments of our new science, and naturalized epistemologists would take a revised stance that we are warranted in believing

---

<sup>18</sup> See Burton Dreben (1994) for a discussion of the Quinean starting point.

those theories that explain and predict our total sensory *and* extra-sensory experience of the physical *and* non-physical world.

Although Stroud is a sympathetic critic of philosophical naturalism, he argues not only that naturalized epistemology fails to address traditional skepticism about the external world, but that it gives rise to a new skeptical argument. The apparent strength of the position against skeptical attacks, he claims, is its removal of the avowedly “philosophical” ground upon which the skeptic presents her challenge that we lack justified beliefs. He quotes Quine:

I am not accusing the skeptic of begging the question; he is quite within his rights in assuming science in order to refute science; this, if carried out, would be a straightforward argument by *reductio ad absurdum*. I am only making the point that skeptical doubts are scientific doubts. (NNK, 258)

But Stroud argues that the skeptic *can* carry out Quine’s imagined *reductio*. According to Stroud, by assuming the veracity of our scientific theory, the skeptic can show that we lack knowledge of the external world since our scientific theory tells us that it is underdetermined by the available data:

The *reductio ad absurdum* would presumably run something like this. Either science is true and gives us knowledge or it does not. If it is not true, nothing we believe about the physical world amounts to knowledge. But if it does give us knowledge, we can see from what it tells us about the meagre impacts at our sensory surfaces during perception that we can never tell whether the external world really is the way we perceive it to be. But if that is so, we can know nothing about the external world. So once again nothing we believe about the physical world amounts to knowledge. (1984, 228)

Stroud’s point is that the skeptic need not assert that she has a *superior* theory to the scientist, but only note that

none of the competing ‘hypotheses’ about what is true beyond the data can be known to be true...[in fact,] we can have no more reason for believing any one of them rather than others on the basis of the only sensory data we can ever have. (1984, 233)

Quine had written that doubting scientific realism on the basis of skeptical reflection was “overreacting” (RS, 475). But Stroud thinks this is glib. The traditional skeptic does not doubt that our scientific theories are the best that we have for *pragmatic* reasons. Rather, she is

disturbed by the realization that *no* scientific theory we can develop amounts to knowledge. On Stroud's analysis, Quine misses the point of traditional skepticism, and naturalized epistemology is a change of subject rather than an "enlightened" development of epistemology.

But it is Stroud who fails to appreciate that Quine is not providing a blanket response to all epistemological challenges that we may call "skeptical." Rather, he diagnoses each skeptical challenge naturalistically, finding some to be answerable and others unintelligible. It is this nuanced attitude that underwrites his charge of "overreacting," and other passages where he gives the skeptic such short shrift:

My answer to skepticism is that reality itself, the term 'reality', the term 'real', is a scientific term on a par with 'table' 'chair' 'electron' 'neutrino' 'class', that all these are part of our scientific apparatus, our terminology, so that the only sense I can make of skepticism is the kind of sense that maybe our theory is wrong, that in the future the checkpoints aren't going to bear it out, checkpoints in observation, in experiment. (EDQD, 152)

This response would also strike Stroud as petulant. How can Quine profess not to *understand* the skeptic's anxiety, which is obviously not captured by noting that our theory could be wrong?

Quine is not being deliberately obtuse. First, he tries to respectfully address the skeptic as a fellow inquirer, engaging a form of skepticism I shall call *constructive* skepticism. He encourages the constructive skeptic to describe the case worrying her in more detail so that she may bring it to the highest tribunal countenanced by naturalists: the scientific tribunal. There, she may advance it as an alternative hypothesis to scientific realism. At this tribunal, the currently endorsed principles of theoretical construction hold sway (simplicity, modesty, conservatism, etc.<sup>19</sup>), and Quine is confident that scientific realism will win out. The *reductio* Quine was actually imagining in the quote Stroud extracts was the (in Quine's view unlikely)

---

<sup>19</sup> Quine sketches these and other desiderata for scientific hypotheses in the sixth chapter of WB.

possibility that a constructive skeptic could describe a theory which is simpler, more modest, and more conservative than scientific realism, which nevertheless enjoyed its explanatory power.<sup>20</sup>

In contrast, Quine would view Stroud's *reductio* as a legitimate expression of fallibilism. He *agrees* with Stroud's skeptic—an inquirer I shall call the *pessimistic* skeptic—that we may come to judge our theory inaccurate. Although scientific realism is presently esteemed as our best theory for predicting sensory experience, reflecting upon the history of science suggests it is highly likely that parts of it are mistaken. We have reason to expect our scientific progression (along with the sloughing off of theoretical blunders and excesses) to continue indefinitely. Accepting that science is fallible, then, is an attitude toward the discipline that the discipline itself—in endorsing inductive reasoning—demands. It is the pessimistic skeptic who withholds her belief in her current theory because of its potential inaccuracy who is, in Quine's view, overreacting, because she has misunderstood the self-acknowledged fallibility of inquiry.

Quine is well aware that this response would seem glib to a third inquirer, who I shall call the *radical* skeptic. The radical skeptic insists that reality might not just *differ* from our current theory, but be *so* different that neither we, nor any future scientists, could come to know it. Yet in Quine's view, this radical skeptic has made a *semantic* error. She fails to articulate an intelligible case. Quine believes, on the basis of his naturalist reflection upon language use and acquisition, that our words only become meaningful in the context of successful ongoing activity with others in the world, language being a “social art.” But by talking about a “reality” that is *in principle* “beyond” our experience, and which exceeds our ability to develop theories, the radical skeptic unbuckles her words from the circumstances in which she acquires them. The problem is

---

<sup>20</sup> A variation is a constructive skeptic who doubts one or more of the tribunal's standards—claiming, for example, that complex theories are preferable to simple ones. In Quine's view, the onus is upon this skeptic to make the best case she can for complexity. Although he doubts that she will succeed, accepting that she *may* succeed is part of his naturalistic attitude. The standards we use to judge theories are also potential candidates for judgment.

that in detaching her words from their semantic mooring, they become meaningless.<sup>21</sup> The radical skeptic warrants no further attention because she cannot coherently phrase her challenge naturalistically.<sup>22</sup>

Grasping this argument brings out the force of Quine's semantic holism. In his view, our scientific theories about what there truly is cannot be strictly distinguished from the meaningfulness of sentences in ordinary language.<sup>23</sup> The radical skeptic's challenge is premised upon "reality" being meaningfully applicable outside the bounds of possible scientific inquiry, but:

naturalism looks only to natural science...for an account of what there is....to ask what reality is *really* like, however, apart from human categories, is self-stultifying. It is like asking how long the Nile really is, apart from parochial matters of miles or meters. (SN, 405)

In Quine's view, "real" is a predicate of our language that applies to those entities our current theory demands. We refine our grasp of what is "real" through scientific activity. Analyzed naturalistically, the radical skeptic's challenge becomes "why am I warranted in believing that reality (which is a term of my theory that I use to refer to whatever I am warranted in believing) is as my current theory tells me (or future theory will tell me) it is?" This rephrasing demonstrates that in seriously asking the question she is misusing the word "reality."

Stroud might object that this defense turns on a claim about the meaning of "reality." He thinks that it is "one of the merits of Quine's views about language that they do not support such

---

<sup>21</sup> Quine writes, "Our words have no meaning beyond what they acquire through our learning of them, and all our learning of them goes back directly or indirectly to the association of utterances with concurrent sensory stimulation...the existence of external objects is itself just one among the tenets of our scientific theory, albeit a primordial one, and it is sustained to the degree that the theory as a whole conforms to observational data. The very meaning of the existence thesis lies no deeper" (FHQP, 206-207).

<sup>22</sup> In another formulation of naturalism, Quine writes that it means accepting that "it is within science, and not some prior philosophy, *that reality is to be identified and described*" (TTPT, 21, my emphasis).

<sup>23</sup> Quine writes, "We learn truth conditions of some sentences relative to other sentences. We learn thus to use the component words to form new sentences whose relative truth conditions are derivable. Which of these dependencies of truth value are due to meaning, or language, and which belong rather to a substantive theory that is widely shared, is in my view a wholly unclear question. It is no mere vagueness of terminology that makes language and theory indistinguishable in this connection" (R, 310).

dubious argumentation,” because, in rejecting “the philosophical use of synonymy or analyticity,” “[Quine is] in no position to appeal to what is or is not included in the meaning of a particular term” (1984, 227). But this objection mistakes the defense. Although Quine’s argument for the indeterminacy of translation precludes him talking about *the meaning* of a term or a sentence in our language, it does not preclude reflection on *the meaningfulness* of (parts of) our discourse. In *Word and Object* and later writings, Quine gives a detailed theory of how we develop our capacity for language.<sup>24</sup> By reflecting on our linguistic practice in this way, he naturalizes the ingredients of traditional epistemology and disarms its problems. The naturalized epistemologist’s response to radical skepticism does not turn upon *a priori* conceptual analysis of the term “reality,” but on a theory of how the ingredient “reality” functions in our language.<sup>25</sup> Quine argues that our understanding of the word “reality” comes from *within* our scientific practice as the name of the goal of our fallible, revisable, theoretical constructions. We acquire competency in its use as we earn membership into the community of inquirers. The word “reality,” like any scientific term, is meaningless outside the scientific practice that makes scientific terms meaningful.<sup>26</sup>

---

<sup>24</sup> I will have more to say about Quine’s theory of language in section three, and about the indeterminacy of translation and the inscrutability of reference in section four.

<sup>25</sup> Stroud might retort that in appealing to one more part of his naturalist *theory*, Quine is relying upon something he *believes* which may thus be doubted by the radical skeptical (1984, 247). But Quine’s naturalist machinery kicks in once again. To the naturalist, *constructive* doubts combined with competing hypotheses are perfectly legitimate, and hence the onus is upon Stroud to give an alternate theory of language—more explanatory, conservative, and modest than Quine’s—which explains how and why humans should have, oddly, developed the word “reality” to refer to something that is forever beyond human experience. If Stroud has no such theory but is merely pointing out, *pessimistically*, that Quine’s could be wrong, Quine will happily admit to fallibilism about his theory of language. If Stroud complains, *radically*, that Quine nevertheless lacks any reason to believe that his current theory of language (or any future refinement thereof) is accurate, Quine will rebuke Stroud for imagining “accuracy” to be a standard for evaluating our theories that is external to those theories.

<sup>26</sup> If we liked, we could of course stipulate a new use for the word “reality.” To borrow Wittgenstein’s terminology, the word might be employed as part of a language game in which we never know how things *really* are. But this language game has nothing to do with the language game of inquiry, and if the radical skeptic is phrasing her challenge using this word, she may justly be ignored by inquirers (though perhaps not by those non-naturalized philosophers who wish to play her game).

A parallel response is available to Quine if the radical skeptic ascends to the formal mode. Objecting that my various beliefs about reality may all be false presupposes a standard independent of all my beliefs, against which they can be compared and found wrong. But Quine denies the existence of any such standard, because there is no external standpoint that we could occupy to apply it.<sup>27</sup> We talk of “truth,” and of aiming to only believe “truths,” to describe the purpose of our inferential practice *from within that practice*.<sup>28</sup> So, Quine can freely admit that our scientific testing procedures—even those as basic as observation—are fallible, and so, may mislead us, without committing himself to skepticism. This is what he means by writing that “the Humean predicament is the human predicament” (EN, 72); although we lack *certainty* about the nature of the world, we can nevertheless successfully develop theories about it (and, moreover, develop theories about our very ability to successfully develop theories about it).

But although this explains how the naturalist can defuse traditional skepticism, Stroud also claims that a new form of skepticism affects naturalized epistemology on its own terms. He begins by examining how a naturalist inquirer will judge whether her interlocutor’s theory is justified. The naturalist’s theory will include an account of what evidence the world makes available to her interlocutor, and, by making a second appeal to her theory for an account of what the world is like, she will be able to evaluate the extent to which her interlocutor’s judgments are correct, and how far her interlocutor’s theory is grounded in her evidence. But Stroud complains that the naturalist cannot follow this procedure to justify her *own* beliefs. Since she would have to appeal to her own theory twice—once for an account of reality, and once for an account of how reality is revealing itself to her—she would be deprived of any “independent information

---

<sup>27</sup> This Quinean response to Stroud parallels the response I gave in chapter one, on Frege’s behalf, to Peter Sullivan. By maintaining that our inferential practice’s standards of evaluation are the *only* standards of evaluation for our beliefs, Frege and Quine inoculate their views from objections that rely upon purported practice-independent evaluative standards.

<sup>28</sup> I will return to Quine’s conception of truth, particularly as it compares to Davidson’s, in chapter four.



about [the] world that [she] could use as a test or check” (1984, 244) for her beliefs, and so, would be incapable of providing any justification *of* those beliefs.

According to Stroud, this amounts to a new skeptical argument:

I could not see my efforts [at justification] as providing me with an explanation that itself is something I know or have reason to believe, as opposed to a more complicated story I fully accept and find myself disposed to tell myself from time to time.” (247)

Yet “justification,” like “truth” and “reality,” is an epistemic ingredient that Quine treats naturalistically. Stroud’s new skeptical argument does not affect naturalized epistemology on its own terms because, for Quine, once we have recognized that there are no standards of evaluation external to our practice of inquiry, we ought to grasp that the only justification that inquirers can have for their beliefs are others in their webs. This should not launch naturalists into skeptical doubt that they forever lack “reason” for their beliefs. Rather, it should encourage them to revise the outmoded philosophical terminology of reasons and justification to better comport with scientific practice. True beliefs are not eternally justified or unjustified, reasonable or unreasonable; they are more or less justified and more or less reasonable, at historical moments.<sup>29</sup> The “reasonability” of our theory as a whole is to be judged in terms of the extent to which it meets the standards of the scientific tribunal, because its standards are the best we have so far discovered for epistemic justification.

In Quine’s view, once we reorient ourselves naturalistically, we see that there is no room for Stroud’s lingering “philosophical” concern that we enter into a vicious circle when appealing to science to explain our knowledge if we have yet to *justify* our knowledge *of* science (EN, 75-76). Rather, we grasp that science and epistemology reciprocally contain each other (RG, 684).

The naturalist uses her best current theory to establish the foundation of her epistemology

---

<sup>29</sup> After proposing this treatment for “justification” under the entry of “knowledge” in his “intermittently philosophical dictionary,” Quine writes: “Here we see the familiar and widely applicable rectification of vagueness: disclaim the vague positive and cleave to the precise comparative” (Q, 109). Rather than adopting the artificial “more or less known,” Quine is content for the word “knowledge” itself to—ironically—go by the board.

(holding that she knows what she knows on the basis of her senses) and reconstructs her ontology using evidence acquired via her senses.

To illustrate this point, Quine frequently turns to Neurath's metaphor (e.g. SO, 16; NK, 127; R2, 178): our epistemic theory is like a boat. We cannot dismantle it in dry-dock and reconstruct it to conform to our preferred design, because we are already at sea. Naturalized philosophers are those enlightened souls who recognize that they stand on the deck with scientists. The crew can work together to repair damaged parts of the boat (and even improve its overall design), but only while keeping it seaworthy, else they risk drowning in a sea of incoherence. Empiricism is a structurally central, though conceivably removable, plank of the boat; naturalism is not a plank at all, but the shared attitude of the crew.

In this section, I have argued that by naturalizing epistemology Quine excises skeptical arguments about the external world. He does so by emphasizing the roles that our language and our evolving scientific theory play in distinguishing meaningful possibilities from unintelligible gibberish. "Reality" and "truth" are internal standards to scientific practice. There is no position of "cosmic exile" to occupy that would allow us to evaluate the extent to which our theories meet putative external standards. In fact, since naturalists deny the existence of external standards, the internal/external contrast that I am employing breaks down. There is *only* the sideways evaluation of our beliefs from a position within our scientific theory. We judge ourselves "justified" according to our scientific standards for theory construction, which are themselves fallible and revisable. In the next section, I shall examine Quine's naturalized conception of language in more detail to show that it excises another skeptical challenge from philosophy.

### **3. Rudolf Carnap and Skepticism about Successful Communication**

[T]he notion of there being a fixed, explicable, and as yet unexplained meaning in the speaker's mind is gratuitous. (WO, 160)

Quine dedicates *Word and Object* to Rudolf Carnap, his “Teacher and Friend.” In this section, I will argue that a skeptical threat facing Carnap’s project in *The Logical Syntax of Language* is excised by Quine’s naturalist approach to language use and acquisition. From Quine’s perspective, Carnap unintentionally invites skepticism about communication by failing to give up first philosophy from his otherwise laudable empiricism. Instead of looking naturalistically at how scientists use language, and then codifying their practice, Carnap imposes structure upon the formal languages he designs for science under the name of “explication.”

The underlying motivation for Carnap’s syntactic project is the promotion of communication. Early in his career, Carnap was frustrated by the impasses that occurred when disputants failed to agree upon the fundamental terms of their disputes, and so, how to resolve them. Such impasses inhibit communal endeavors. Carnap tried to adopt his current interlocutor’s chosen vocabulary, switching vocabularies between conversations as necessary; but this approach seemed disingenuous to others.<sup>30</sup> *The Logical Syntax of Language* is Carnap’s defense of his practice.

Carnap details a syntactic general method for constructing formal languages. A consequence relation can be clearly defined for each language, fixing which of its expressions imply which others (LSL, 168).<sup>31</sup> Carnap’s idea is that participants in a recalcitrant dispute can

---

<sup>30</sup> Carnap writes, “I [became] aware that in talks with my various friends I had used different philosophical languages, adapting myself to their ways of thinking and speaking. With one friend I might talk in a language that could be characterized as realistic or even as materialistic...with another friend, I might adapt myself to his idealistic kind of language...I was surprised to find that this variety in my way of speaking appeared to some as objectionable and even inconsistent” (IA, 17). See the final chapter of Michael Friedman (1999) for an illuminating account of the evolution of Carnap’s attempts to approach disputes evenhandedly.

<sup>31</sup> To fill in a few details, *formation rules* specify which strings of symbols count as expressions of the language, and *transformation rules* specify which expressions of the language are consequences of others.

agree to *defer* their disagreement until each has endorsed the consequence relation given in a syntactically identified formal language. After their syntactic digression, if disputants endorse *the same* formal language, then they can return to their original dispute, each confident that they know what the other will count as a good argument. They may appeal to the formal language as needed, investing its meaningless marks with meanings via an “interpretation,” and determining to their mutual satisfaction whether to endorse or reject any disputed claim on the basis of its testable consequences (LSL, 227-233; FLM, 37).<sup>32</sup> Where disputed terminology is vague, inquirers can use the natural language they share to informally describe the features that interest them (the *explicandum*), and then define a corresponding expression in their chosen formal language which preserves those features (the *explicatum*):

An explication replaces the imprecise explicandum by a more precise explicatum. Therefore, *whenever greater precision in communication is desired*, it will be advisable to use the explicatum instead of the explicandum. (SLM, 935, my emphasis)

Since picking a formal language makes explicit the consequence relation to which inquirers will hold each other accountable, it also promotes “precision in communication.” Using Carnap’s machinery is meant to reassure inquirers that their dispute is, in principle, resolvable. There is a mutually binding standard for their arguments, against which they count as agreeing or disagreeing about precise claims rather than merely talking past each other.

In contrast, disputants who endorse *different* formal languages are only engaged in *pseudo-dispute*. Their putative “disagreement” seemed recalcitrant because they do not recognize the same arguments as binding. Their discussion lacks a shared standard against which they count as disagreeing. If they wish to continue collaborating, one must persuade the

---

<sup>32</sup> Once a formal language is interpreted, logical vocabulary can be distinguished from physical vocabulary, and consequences the inquirers agree to be logical can be distinguished from those they take to be physical (LSL, 177; 181).

other to endorse her preferred consequence relation. Carnap is unreceptive to interlocutors who refuse this method:

If one partner in a philosophical discussion cannot or will not give a translation of his thesis into the formal mode, or if he will not state to which language-system his thesis refers, then the other will be well-advised to refuse the debate, because the thesis of his opponent is incomplete, and discussion would lead to nothing but empty wrangling. (PLS, 80-81)

Since the purpose of his syntactic method is to promote communication, those who decline it are justly refused.<sup>33</sup>

Carnap thinks that the persuasion between collaborators about which consequence relation to endorse is purely pragmatic, which is the sense of his slogan, “*In logic, there are no morals*” (original emphasis, LSL, 52). Arguments that inquirers form “internally” within a language derive their force from the language’s mutually agreed upon consequence relation. From within, one inquirer can *show* another wrong. But arguments formed “externally” lack the force to bind disputants.<sup>34</sup> There is no precise application for “right” or “wrong” prior to agreement about formal language. So in choosing a formal language, inquirers rely solely on its perceived adequacy to their current endeavor and may only try to persuade others of its expediency. If two inquirers persist in holding different languages to be satisfactory—in Carnap’s most developed example, classical and constructive mathematicians—then a metaphilosophical attitude of tolerance between them is appropriate: “it is not our business to set up prohibitions, but to arrive at conventions” (LSL, 53). In this way, disputants who refuse to speak the same language are to part ways amicably.<sup>35</sup>

---

<sup>33</sup> It perhaps goes without saying that many metaphysicians were unhappy being refused or “tolerated” by the logical positivists, rather than engaged—voicing a frustration of the merely tolerated everywhere (see e.g. Gustav Bergmann, 1954, 64-65; Philipp Frank, 1963, 159-164).

<sup>34</sup> Carnap argues that the complexity of natural languages thwarts the formal treatment necessary to clarify their consequence relations (LSL, 2; TLS, 50). Since it cannot be assumed that speakers of the same natural language understand consequence identically, arguments about formal language choice that are phrased in natural language are external and only have the potential to compel pragmatically.

<sup>35</sup> By making the set of truths—and in particular, the set of *logical* truths—that each inquirer endorses a function of her pragmatic choice of formal language, rather than taking all inquirers to be constitutively committed to a single,

I have been arguing that Carnap's motive is to promote communication. But the skeptical challenge that I have in mind questions whether those using his machinery can know that they *are* communicating. What sort of guarantee do disputants have that they are not talking past each other once they have pragmatically chosen a formal language?<sup>36</sup> Carnap says that inquirers can resolve internal disputes by "interpreting" the meaningless marks of their formal language and determining the testable consequences of disputed claims. But how do inquirers know that they are interpreting the formal language in the same way?

We can consider Carnap's reply to a related objection from E.W. Beth to reconstruct the response that he would give this skeptic. Beth charges Carnap with having made an illicit assumption about his reader. He argues that one must tacitly assume one's intuitive interpretation of mathematics to understand Carnap's syntactic description of the meaningless marks of different calculi, and in particular those marks which allow for the construction of arithmetic within the calculus (once interpreted). Carnap's illicit assumption is thinking his

---

universal logic that "unfolds" the concept of truth, Carnap's logical pluralism is an important step away from Frege's universalism. The tolerance of a Carnapian might even extend to the second species of logical aliens. She may find logically alien inferences baffling, and lack any pragmatic reason to adopt their seemingly contradictory formal language; this need not be a reason to deny them cognition, as opposed to tolerating their puzzling practice from a safe distance.

<sup>36</sup> This question is pressing on the interpretation that I am pursuing, but Friedman has offered a different characterization of the aim of Carnap's syntactic method:

Carnap is not worried about determining, in actual cases, which disputes are genuine and which are not. He is already perfectly clear about this: *philosophical* disputes are characteristically fruitless, whereas *scientific* questions...patently are rationally negotiable. Carnap's problem is not to discriminate the fruitless disputes from the fruitful ones but to offer those enmeshed in the former—philosophers—a way out. This is what the construction and investigation of a variety of formal languages is for (1999, 214 n29, original emphasis).

According to Friedman, Carnap acknowledges that his syntactic method will only help "scientifically minded, nonmetaphysical" philosophers to communicate (213). He does not need to provide such inquirers a skeptic-proof guarantee that they *are* communicating once they employ his syntactic method, because they will have no interest in such philosophical questions. In my view, Friedman's interpretation undercuts the force of Carnap's refusal to engage those who reject his syntactic method. Carnap's advice that inquirers avoid "empty wrangling" looks less like a well-intentioned recommendation to avoid obscurantists who refuse to clarify the terms being debated, and more like a prejudice against philosophy. But even if the argument that I develop here does not undermine Carnap's aim, I believe it demonstrates that, unlike Quine, he is forced into a position that is susceptible to criticism by "metaphysically minded" philosophers.

readers will understand the calculi as he intends them to, in accordance with the interpretation that seems intuitive to *him*.<sup>37</sup> But Beth argues that if the model of arithmetic the reader finds intuitive is non-standard, she will be unable to understand or use Carnap's syntactic method. On page 113 of Carnap's book, Beth claims, she will judge the definition of "analytical in II" inadmissible and "will no longer be able to follow [his] argument" (1963, 481).<sup>38</sup>

Carnap replies by urging Beth to carefully distinguish "models" from "interpretations." He defines a *model* as an assignment of extensions to all the primitive vocabulary in a language and an *interpretation* as an assignment of meanings to all the signs of the language.<sup>39</sup> Any reader with a non-standard model of arithmetic must also have a deviant interpretation of our numbers. But for Beth's objection to go through, such a reader must also interpret the metalanguage (ML) in which *The Logical Syntax of Language* is written (and in particular its claims about numbers) in accordance with her deviant interpretation. Hence, Beth surprises Carnap, who writes:

I always presupposed...that a fixed interpretation of ML, which is shared by all participants, is given. This interpretation is usually not formulated explicitly; but since ML uses English words, it is assumed that these words are understood in their ordinary senses. *The necessity of this presupposition of a common interpreted metalanguage seems to me obvious.* (BCS, 929-930, my emphasis)

---

<sup>37</sup> Beth insists that "it is completely consistent with current word usage to denote [a reader's] way of understanding Carnap's discussion as an interpretation" (1963, 483) and, furthermore, shows that such usage fits Carnap's own usage of "interpretation" as a translation (LSL, 228). Beth thinks that each reader interprets *Syntax* by "translating" it into her own intuitive understanding of the topics it discusses.

<sup>38</sup> I am omitting the technical details of Beth's argument and how it relates to the Lowenheim-Skolem theorem. For critical discussion, see Thomas Ricketts and Warren Goldfarb (1992), Ebbs (1997), and Michael Friedman (1999).

<sup>39</sup> Carnap writes: "an interpretation should not be identified with a model, as is sometimes done. It is true that an interpretation can be given by the specification of a model. But there is not a one-one correspondence between interpretations and models; two different (i.e., not logically equivalent) descriptions of the *same* model represent two *different* interpretations" (MCS, 902, original emphasis).

Carnap reduces Beth's objection to an "obvious presupposition." Of course, someone might misunderstand his syntactic project, but so might someone misunderstand *any* book or conversation, if they use a deviant interpretation of natural language.<sup>40</sup>

Carnap's acceptance of the "necessity" of a "shared," "fixed interpretation of ML" as a "presupposition" of his project fuels a more serious objection. Although we might identify non-standard mathematicians in our midst by making our set-theoretic assumptions explicit and discovering that theirs differ from ours, there is no reason to think that *every* difference between inquirers' interpretations will reveal itself in some difference between their dispositions to affirm or deny particular sentences. Moreover, *any* language in the infinite hierarchy to which inquirers using the syntactic method might appeal to reassure themselves that they are interpreting expressions in the same way is also subject to questioning, which seems to trap them within an infinite regress.<sup>41</sup> In my view, Beth's objection exposes that Carnap is committed to thinking his project only *promotes* communication and cannot *guarantee* it, an epistemic gap that a skeptic can exploit to question whether we ever *do*, rather than *seem*, to communicate.

To see why, suppose on Monday two physicists, Amy and Ahmed, have a seemingly intractable disagreement. They both read Carnap's book and pick Language II as the most expedient calculus for their goals.<sup>42</sup> They now return to their disagreement, carefully explicate its terms, run tests to determine which consequences obtain, and resolve their dispute. Pleased at a hard day's work, they continue their collaborative investigation. But on Friday, Ahmed uses a word that was part of the explicandum of one of the terms of their earlier disagreement in a way

---

<sup>40</sup> Carnap gives the example of a reader who understands "no occurrence of x" to mean "at least one occurrence of x," an error which would quickly lead to difficulties (BCS, 930).

<sup>41</sup> Although both Ricketts and Goldfarb (1992) and Ebbs (1997) reconstruct the force of the Beth objection as an infinite regress, I differ from them in thinking that it exposes a difference between the responses Quine and Carnap are able to give the skeptic.

<sup>42</sup> Carnap writes that physics can be formulated using a language of the form of II—i.e., one whose transformation rules codify the classical logical laws (LSL, 150).



that Amy finds baffling. Experimenting with each other's linguistic dispositions to various test sentences, they discover a subtle difference in their ML-interpretation of the word that they had each unwittingly carried over into their conversations using the explicatum. The difference is serious enough to undermine their confidence in the "resolution" of their prior dispute. Although each thought that when using the explicated word during the week they *had been* successfully communicating, they both now realize that they had *not* been communicating after all. Each failed to understand the thoughts that the other was intending to express, and so, they were neither agreeing nor disagreeing.<sup>43</sup>

Carnap would accept that Amy and Ahmed have not been communicating all week, but he would encourage the physicists not to be unduly alarmed by this hiccup. There is an "inevitable vagueness" to ordinary language (BCS, 930). Having identified the subtle difference in their usage of the problematic word, they can now disambiguate it. They can then explicate each disambiguation, and go on collaborating.<sup>44</sup>

But now, suppose that Amy and Ahmed are led by this event to seriously question their tacit assumption that they share an interpretation of ML. This particular disagreement was brought to light by the happenstance of Ahmed's using a word in a way to which Amy objected. Now they realize that a dormant disagreement could underlie *any* part of their language. However confident they are that they are communicating, and however productive their mutual projects, it is possible that a discovery tomorrow will reveal their speech dispositions to have been only accidentally coordinated. It is even intelligible that such accidental coordination

---

<sup>43</sup> Carnap took understanding the thoughts others express to be one of the purposes of linguistic communication: "Throughout my life I have been fascinated by the phenomenon of language. How amazing and how gratifying it is that we are capable of communicating with one another by spoken sounds or written marks, *to describe facts or express thoughts and feelings*, to influence the actions of others" (IA, 67, my emphasis).

<sup>44</sup> This advice presumes that Amy and Ahmed did not depend upon the problematic word when arriving at their understanding of the consequence relation of their formal language. If they did, their different usage may reveal that they actually want to endorse different formal languages. In this case, Carnap would advise Amy and Ahmed to begin pragmatically defending their choice of formal language until one concedes, or else to part ways.

happens globally. Amy and Ahmed realize that they could syntactically investigate their ML using a meta-metalanguage (MML) and run empirical tests to ascertain each other's speech dispositions to particular cases; but since they may interpret MML differently, this will not be definitive. Desolate, Amy and Ahmed conclude that they will forever lack certain knowledge that they are communicating, as opposed to merely seeming to.<sup>45</sup>

Ebbs has argued that Carnap might respond to this objection by insisting that inquirers achieve the clearest purchase possible on their inquiry's argumentative structure by explicitly choosing a formal language (1997, 124-125). Syntactic investigation is the best method available for inquirers to certify that they are agreeing. Although it is necessary to presuppose that inquirers share a fixed interpretation of the ML to *begin* syntactic investigation, the meaning of "agreement" is only fixed once this investigation has been completed. Amy and Ahmed should recognize the consequence relation of their chosen formal language as the baseline of their communication.

Ebbs also contends that Carnap would think it a "confusion" to talk of an "implicit interpretation of the metalanguages of pure semantics" (1997, 123)—but, as evinced by the above quote from his response to Beth, Carnap himself acknowledges that a "fixed interpretation of ML" is a "necessary presupposition" for applying his syntactic method. In a move resembling Quine's to the radical skeptic, Ebbs would have Carnap denying Amy and Ahmed the ground upon which to phrase their challenge: "there is no higher or firmer perspective from which to question whether investigators 'really' share the metalanguages whose expressions they find clear [than those shared metalanguages]" (124).

---

<sup>45</sup> We may well counsel Amy and Ahmed to be optimistic. After all, they have had only one disagreement thus far, the supposed accidental coordination would be an astonishing fluke, and so on. But Amy and Ahmed need not find any of these considerations compelling. If communication names the purported ability to express one's private thoughts to another mind using the medium of language, then one may intelligibly doubt that it ever succeeds.

But the analogy between Carnap and Quine fails. Without any promise of *certainly* that they are communicating, Amy and Ahmed may skeptically refuse to take Carnap's optimism seriously. Since it is part of Carnap's view that genuine communication between inquirers requires them to share an interpretation of their language, so that each can grasp what the other intends to express, he is forced to acknowledge Amy's and Ahmed's failure to communicate as an intelligible possibility, rendering their skeptical doubt *cogent*.

In contrast, by naturalizing the ingredients of epistemology, Quine renders various radical skeptical positions—including Amy's and Ahmed's doubts—*unintelligible*.<sup>46</sup> He develops his account of language acquisition and use to further his argument against the analytic/synthetic distinction.<sup>47</sup> Prior to *Word and Object*, Quine's positive attack against the distinction focused upon its vagueness. In formally constructing a language, one may (as Carnap does) *label* some truths "analytic," but unless "analyticity" tracks a property that exists outside the language, Quine objects that the label is arbitrary. Analytic truths may be informally described as those expressions that are synonymous with logical truths, but this just shifts the explanatory burden onto the relation of synonymy. One might attempt to define "synonymy" as holding between

---

<sup>46</sup> On the surface, Quine's claim that we overcome skeptical doubts about ontology by "acquiescing in the home language" seems to parallel Carnap's that we must assume "a fixed interpretation of ML" to apply the syntactic method. But Quine's naturalized, extensional understanding of what it is to share a home language removes the intensional vocabulary that Carnap's "interpretation" leaves in the picture. The force of Quine's view is not just that there is no higher or firmer ground from which to question ontology than our home language, but that once naturalists properly understand what questioning ontology amounts to, they can no longer phrase *radical* skeptical doubts about it. Under Carnap's view, Amy and Ahmed can do so. I will return to Quine's ontological relativity in the next section.

<sup>47</sup> One might object that Carnap is closer to Quine than I am suggesting since, if pressed, he would try to explain language use in behaviorist terms. But Carnap's forays into descriptive semantics, as I mention below, are crude, and his proposed behaviorist basis for attributing intensional attitudes is contentious. In fact, he seems to think that "understanding" can be made precise by making a semantic, non-behaviorist appeal to "interpretation" (which, I have been arguing, relies upon a shared understanding of the language in which the interpretation is given). In discussing the electric field vector "*E*" in Maxwell's equations, he writes that "we understand '*E*,' if 'understanding' of an expression, a sentence, or a theory means capability of its use for the description of known facts or the prediction of new facts," once we are given "an interpretation for '*E*'...indirectly by semantical rules referring to elementary signs together with the formulas connecting them with '*E*'" (FLM, 68-69). In future work, I hope to build on the reading of Carnapian explication I offer below to argue that Carnap is unable to provide a full behaviorist explication of "understanding" which avoids relying on "interpretation" and is robust enough to ground his account of analytic truths as those which can be determined true once they are understood (QLT, 916).

any expressions that possess the same “meaning,” or that express the same “proposition,” noting that a dictionary can be used to locate the meaning of an expression and a thesaurus to locate words with synonymous meanings. But Quine responds that dictionaries do not *imbue* words with meanings. They only *report* the meanings that words are taken to have by compilers. Likewise, thesauruses do not ground synonymy, but presuppose it.

For Quine, the problem is not just that “synonymy” is opaque. The problem is that it is scientifically unnecessary—and so naturalistically inadmissible—to posit “meanings” to account for language use. Reference is behaviorally inscrutable; it is logical theory, not language use, that requires precise semantic notions. But the purely extensional notions of “truth” and “reference” suffice for logic, and so—given logic’s canonicity—for scientific inquiry. Thus, Quine advises ejecting a slew of inter-definable, but scientifically mysterious terms from naturalist philosophy. “Meanings,” “intensions,” “necessary,” “analytic,” “synthetic,” and “propositions” are to be removed or replaced with extensional analogues as needed.<sup>48</sup>

To justify this purge, Quine needs to provide an adequate extensional account of language and to explain why our practice of *using* dictionaries and thesauruses does not ground a naturalistically admissible notion of synonymy. For even granting that dictionaries do not *imbue* words with meanings, if language users learn from their early education (supplemented by an occasional glance at a dictionary or a thesaurus) that two terms can be used interchangeably, this

---

<sup>48</sup> Quine justifies his career-long partiality for extensional theories on the basis that he “doubt[s] that [he has] ever fully understood anything that [he] could not explain in extensional language” (500). Concepts that are defined extensionally, by isolating the class of objects of which they are true, can be straightforwardly manipulated in logical derivations. In contrast, concepts that are defined intensionally, as “properties,” are not usually determined by their instances, and so, “lack...clear principle[s] of *individuation*” (500, original emphasis). Allowing intensions has the effect of admitting vagueness into our theories, creating space for systematic confusion between inquirers who tacitly disagree about the “meaning” of their concepts. Quine refuses to accept these consequences, insisting upon the “clarity and convenience conferred by extensionality” (498). The sole exception is Quine’s acceptance of Davidson’s anomalous monism. Quine allows the ascription to others of propositional attitudes, since their “applicability is outwardly observable enough for practical utility” (quoted in Dreben 2004, 291). Nevertheless, Quine views the ascription of mental “entities” to others lying behind these attitudes, in anything stronger than the physicalistic terms of hypothetical states of their nervous system, as scientifically illegitimate (R, 296).

fact will be reliably reflected in their observable speech behavior. So what prevents the behaviorist (and naturalistically acceptable) definition of two expressions being synonymous if speakers possess identical dispositions to use them in speech?

In *Word and Object*, Quine aims to meet both obligations by arguing for the indeterminacy of translation. He argues that if two people are disposed to use an expression in the same way, it does not entail that that expression has the same meaning for them.<sup>49</sup> There are numerous equally acceptable but mutually incompatible ways to translate another's speech behavior. There is no fact of the matter about the correct translation of another's utterance, and likewise, no fact of the matter about *the meaning* of an expression in natural language. Consequently—and surprisingly—"synonymy" is unnecessary to explain the success of translation.

Despite first appearances, the indeterminacy of translation does not undermine language use. But it does question folk wisdom about what constitutes successful communication. We saw in section one that Quine's radical translator approaches the natives' linguistic behavior as a phenomenon to be explained. One explanation of the usefulness of such behavior to the natives will be provided if she develops a theory of translation for it in terms of her own language. Since the language is unlike any she knows, she does not bring any pre-theoretic assumptions.<sup>50</sup> Her experimental approach is naturalistically admirable. Quine thinks philosophers of language should similarly leave behind their pre-theoretic assumptions. If we consider a human language naturalistically, it is just a system of noises that some portion of the population register as

---

<sup>49</sup> Nor will an individual be disposed to use two expressions thought synonymous in just the same way: for example, if asked for a single word referring to a man who has never married, a person will be disposed to answer "bachelor" and not "unmarried man" (WO, 46).

<sup>50</sup> Even her tacit hypothesis that the natives *are* speaking a language is revisable. If she can find no entering wedge into their utterance patterns, and sustained observation reveals that the native sounds merely accompany and do not further native endeavors, she may find an alternative explanation for the phenomenon: the natives are not language users after all.

meaningful. But “meaningfulness” need not be cashed out in terms of individual noises possessing individual meanings. Posits like “meanings” are only legitimate if they are demanded by our ongoing theory, striving as it does to be economical. We must keep in view that the basic phenomenon a naturalist theory of human language seeks to explain (despite the intensional rhetoric of working psycholinguists) is the genesis and continuation of our system of meaningful noises: how is it that each of us can listen to noises other speakers make and respond with noises of our own, in a coordinated way that furthers shared projects?<sup>51</sup>

This marks a crucial difference between Carnap and Quine. Quine had objected that Carnap’s syntactic project was either a method for *establishing* the logical truths in syntactically given languages, or a method for *codifying* the logical truths recognized in natural languages already in use (TC, 97-99). But if the former it was question-begging, since “logic is needed for inferring logic from the conventions” (97). And if the latter, Carnap needed to supplement his theory with an account of how the formal language a person used could be read from her behavior. Yet we have seen in contrast that Carnap conceived of his project as *explicative*. He thought speakers of a natural language could use a formal language to *replace* their unclear discussions whenever necessary. There need not be clear correlates in natural language for the precise formal terms in which inquirers carry out their discussion (indeed, if their natural language had such precise terms, their dispute would probably have been resolvable without recourse to the syntactic method).<sup>52</sup> In endorsing a formal language, inquirers are not “inferring”

---

<sup>51</sup> Quine’s naturalist attitude to phenomena requiring explanation even extends to his *self*-understanding. *He* is a phenomenon that science can explain: “Light rays strike my retinas; molecules bombard my eardrums and fingertips. I strike back, emanating concentric air waves. These waves take the form of a torrent of discourse...my ability to strike back in this elaborate way consists in my having assimilated a good part of the culture of my community...all this training consisted in turn of an impinging of physical forces, largely other people’s utterances, upon my surface, and of gradual changes in my own constitution consequent upon these physical forces” (SLS, 215).

<sup>52</sup> Carnap writes of explications that “[t]he explicatum [might] not belong to the ordinary language originally, but [be] introduced as a scientific term” (SLM, 935-936).

logical truth by “establishing” the conventions they will adopt. Nor are they “codifying” the logical truths to which they are committed in natural language, so that their speech behavior acts as a criterion of correctness for putative codifications. Instead, their choice of calculus explicates and replaces whatever prior, vague conception of logical truth they had.<sup>53</sup>

Quine rejects this “explicative” project as disguised first philosophy. Although he praises Carnap for recognizing that one should not judge a putative explication by evaluating whether its explicatum captures the *meaning* of the explicandum but, rather, by whether its explicatum can be *used* instead of the explicandum (WO, 259), their understanding of explication, and of “usefulness” as a criterion for judging them importantly differs.

Quine’s paradigm for explication is the ordered pair. Inquirers saw the value of treating pairs of objects as single objects in derivations, but were strictly speaking unable to do so (since the noun “ordered pair” failed to refer to an object, and so, an ordered pair could not occur as the value of a first order variable). Although various formally legitimate definitions for the noun were proposed, Kazimierz Kuratowski’s was eventually adopted, not as the *correct* or *true* definition of “ordered pair,” but as the one most *helpful* for subsequent inquiry.<sup>54</sup> Quine thus evaluates the usefulness of an explication in terms of its role in enabling desired derivations within canonical notation.

Carnap has a more liberal account of explication. He treats replacements that purport to sharpen the boundaries of *concepts* as explications, as when he claims “piscis” as the scientific explication of “fish” (LFP, 6). Such explications are useful, he argues, because concepts whose

---

<sup>53</sup> As Ricketts puts it, for Carnap “linguistic behavior is, so to speak, in itself logically amorphous” (2004, 193). Inquirers only fix the logic of their ordinary discourse by explicitly choosing a formal language.

<sup>54</sup> Kuratowski defines the ordered pair  $\langle x, y \rangle$  as the set  $\{\{x\}, \{x, y\}\}$  (1921). This definition has the virtue of remaining within the domain of the ordered pair’s elements.

correct application are clear in some cases but unclear in others become replaced with concepts whose correct applications are always clear (7).<sup>55</sup>

That Quine's paradigm case for explication concerns a noun (while Carnap explicates concepts) does not merely display Quine's preference for the extensional. Rather, Quine's naturalism imposes a restriction upon explication. In his view, explication is only legitimate when inquirers have discovered a "defective noun" (257) that they wish to use in derivations, but for which they have yet to develop the needed formal expression.<sup>56</sup> In this strong sense, scientific inquiry issues the demand for, and imposes the formal constraints upon, the naturalist philosopher's explications. But the motivation for Carnap's proposal to explicate concepts by imposing conventions about their correct use is less clear, and, in some cases, doing so allows extraneous philosophical assumptions and distinctions—such as the core idea in Carnap's syntactic method that some truths are "analytic"—to infect scientific theory.<sup>57</sup> As Daniel Isaacson astutely observes, "Whereas Carnap sees formal syntax as constitutive of meaning, for Quine a formal system must be a *formalization* of some *already* meaningful discourse" (1992,

---

<sup>55</sup> Joseph Hanna credits Abraham Kaplan with distinguishing between *internally vague* terms that require us to "[find or construct] objects which serve the desired functions," and so warrant Quinean explication, and *externally vague* terms which neither obviously include nor exclude certain objects, and so warrant Carnapian explication (1968, 35). Hanna argues that Quine and Carnap both failed to distinguish between these two types of explication.

<sup>56</sup> Quine writes that useful nouns which fail to refer are also defective, such as "frictionless surface" (WO, 249). He dismisses these "as irreferential fragment[s] of a few containing phrases" (257).

<sup>57</sup> Richard Creath has recently suggested that Quine's late acceptance of "externalized analyticity" as applying to those "sentence[s]...[one] learn[s] to assent to...by learning one or more of [their] words" (PT, 55) constitutes an empirical criterion for logical truth, thus rendering legitimate Carnap's proposal for explicating analyticity in terms of those sentences "true in virtue of meaning" (2004b, 59-60). Creath's idea is that since Quine finds *some* analytic/synthetic distinction intelligible, he can no longer object that Carnap's preferred explication of the distinction is unintelligible in virtue of it having no clear explicandum in natural language. But Quine is explicit that his criterion "gives no clue to the demarcation between analytic and synthetic sentences that has exercised philosophers, out beyond where anyone either remembers or cares how he learned the pertinent words" (PT, 55); he is introducing "analyticity" as a predicate applying to a subset of a speaker's sentences—those constructed from a core set of words which the speaker can recall learning—in order to do some "justice to the intuitive notion of tautology," "the notion that comes into play when we protest that someone's assertion comes down to "0=0" and is an empty matter of words" (55). Quine remains convinced that "analyticity" lacks an explicandum in natural language. This is the heart of his objection to the distinction. He accepts its heuristic value (RHB, 93-94). But its philosophical utility is insufficient justification for misrepresenting some truths as privileged within one's scientific theory, true "in virtue of meaning," and (in some sense) beyond repute.



125, original emphasis).<sup>58</sup> Carnap, that is, takes scientific theories to eventually be semantically grounded by, and in, the formal languages that his syntactic method allows inquirers to construct. He sees no problem in “explicating” some theory by introducing formal distinctions that sharpen its terms. But Quine inverts the semantic dependence, insisting that the purpose of canonical notation is only to clarify theories which are semantically grounded in scientific practice. He only allows “explications” to *refine* meaningful scientific inquiry and refuses to press upon theories conventions, novel terms, or distinctions that they do not require.

The restriction that naturalist philosophers must only follow the demands arising from scientific practice is the reason that Quine thinks the naturalized philosopher of language stands to learn from the figure of the radical translator, and, in particular, her presuppositionless investigation of native linguistic behavior. Radical translation reveals which assumptions are *needed* to explain the meaningfulness of a language to its users. Quine argues that positing “the meaning” of native sentences like “Gavagai” is not demanded. All that matters is that the manual can be used to further shared projects between translator and natives. For example, being able to reliably garner assent when she utters “Gavagai” to natives in the presence of rabbits will allow them to quickly decide about the choice of prey on hunting trips. In practice, the radical translator would come to favor a translation of “Gavagai” as “rabbit” for simplicity, and may informally express this by saying that the two words are “synonymous,” or have the same “meaning.” But this translation is just her (sensible) preference as a manual-maker and has no claim to being the *actual* translation of the native word. She could just as well have adopted a

---

<sup>58</sup> Incidentally, I conjecture that this difference explains Carnap’s fascination with, and Quine’s total lack of interest in, an “international language” like Esperanto. For Carnap, an artificial universal language was an exciting proposal for international scientific communication (IA, 69). For Quine (who was always interested in learning new *natural* languages, speaking German, French, Portuguese, Spanish and Italian fluently), Esperanto represented a basic mistake (Q, 10). The terms of such a language fail to explicate any precisely agreed upon explicandum (c.f. Carnap’s freewheeling use of “Western culture” as the explicandum for a universal European language [IA, 70]).

manual in which “Gavagai” was translated as “undetached rabbit-part,” since “[t]he only overall test for a good manual of translation [is] fluent dialogue and successful negotiation” (quoted in Dreben 2004, 289).

Quine thinks that the naturalist philosopher of language should acknowledge that these considerations apply generally: “radical translation begins at home” (OR, 46). Like Amy and Ahmed, we typically adopt a simple homophonic translation when talking to each other. If faced with disconfirming behavioral evidence, we reject our homophonic hypothesis and speculate that our interlocutor is using a word or some words differently. But just as the radical translator’s preference for simplicity was no ground for supposing the existence of a “correct” translation of the natives, so too have we no ground for thinking that there is a “meaning” onto which we have latched in our preference for homophonic translation. Homophonic translation is acceptable because it accords with the behavioral evidence we have of each other’s speech dispositions, and it allows us, *ceteris paribus*, to co-ordinate our activities in pursuit of mutually held goals; but other manuals could do this equally well.<sup>59</sup>

Quine concludes that loose intensional talk is naturalistically unprincipled. Despite purporting to share Quine’s behaviorist scruples when investigating languages in use (TM, 454), Carnap’s toy accounts of descriptive semantics are wildly optimistic about the sort of intensional

---

<sup>59</sup> Quine was forced to repeatedly emphasize that the indeterminacy of translation in no way impugns interlinguistic dictionaries which are (legitimately) created to be as simple as possible for convenience (OR 34; PT 48-49; RJW, 728). Nevertheless, commentators still miss Quine’s point. Hans Glock, for example, after emphasizing that Quine must view the translator’s early identification of native assent and dissent as a hypothesis, writes:

if there is no fact of the matter as to whether ‘gavagai’ refers to rabbits, then, by parity of reason, there is *no fact of the matter* as to whether the native assents to or dissents from the translator’s ‘gavagai’. Within Quine’s framework, that would remove the possibility of translating anything. (2003, 180, original emphasis)

But, Quine can accept that the translator’s identification of assent and dissent in the native tongue is an analytic hypothesis that is, strictly speaking, indeterminate while maintaining that translation is possible, since his point is that *any* successful translation manual goes beyond the observable evidence available to the behaviorist empiricist (R, 312). In Quine’s view, as we saw in section one, even the translation of truth functional logic is indeterminate (R, 317-318).

features behaviorists can observe.<sup>60</sup> From Quine's perspective, Carnap's claims that disputants can endorse "mutually understood" formal languages given their "ordinary interpretation" of ML, in which they can agree upon the "meaning" of inviolable "analytic" rules, is first philosophy laced with intensional fiction.<sup>61</sup> Having shown intensional vocabulary extraneous to his naturalized theory of language, Quine denies that "successful communication" is a matter of successfully grasping the "thoughts" one's interlocutor is trying to express. Instead, he provides a behaviorist ersatz for communication as *constituted* by the lack of breakdown in our conversation when we engage in shared projects: "Folk wisdom has it that we communicate successfully because our sentences *mean* alike for us...we now see rather that meaning alike for us merely means, if anything, that we are communicating successfully" (quoted in Dreben 1992, 305).

Quine's conception of communication excises Amy's and Ahmed's skeptical doubt.

Carnap's syntactic method allowed them to worry that *despite* their apparent agreements and

---

<sup>60</sup> Carnap imagines "[a] certain lake in [the native] country, which has no name in English, is usually called 'titisee'. When using this name, the people *often think of plenty of fish and good meals*" (FLM, 5, my emphasis). Carnap leaves mysterious what entitles the behaviorist to describe the native's thoughts.

<sup>61</sup> Carnap makes the claim that speakers of a natural language share a more or less fixed "ordinary" interpretation later in his response to Beth. Beth had described the *strict usage* of natural language as uses where speakers "refer to a definite model of the theory to which [their] statements belong" (1963, 479). Perceiving vagueness in this description, Carnap explicates it:

I assume that the strict usage of ordinary language is meant to be *the usage based on a fixed interpretation*, presumably in such a way that the words are understood in their *ordinary meanings*...if my above interpretation of the term "strict usage" is correct, then Beth's thesis says that it is essential for the purposes of my theory that the English words of my metalanguage ML are sometimes used with a fixed interpretation. *I emphatically agree; I would even say that this is the case not only sometimes but practically always.*" (BCS, 930, my emphasis).

Carnap thinks it necessary for his readers to share his "model" and "interpretation" of ML if they are to understand him. But Quine's thesis of ontological relativity entails that "it is...*meaningless* within the theory to say which of the various possible models of our theory form *our real* or intended model" (OR, 54, my emphasis). Quine does not say that it is *difficult* to identify or specify our real model, given the "inevitable vagueness and ambiguity of ordinary words" that Carnap himself goes on to mention in his response to Beth (BCS, 930). Rather, it is *meaningless* to say that our theory *has* a "real" or "intended" model, because ontology is intelligible only with respect to a manual for translation. For Quine, Beth's notion that there is such a thing as "strict usage" of natural language is more intensional fiction.

disagreements, they were nevertheless failing to communicate in virtue of interpreting ML differently. But once the physicists reorient themselves naturalistically, they ought to realize that the case worrying them—that they are only seeming to communicate while actually talking past each other—is not intelligible. Communicating just *is* seeming to communicate in virtue of engaging in productive shared projects and being able to predict how each other will vocally react to stimuli in the environment.<sup>62</sup> If their shared project falters on Friday, Amy and Ahmed ought not conclude, with Carnap, that the hiccup reveals they have not been communicating all week (but can now use his syntactic method to try once again to do so). Instead, with Quine, they should revise their translation manuals for each other so as to *continue* successfully communicating.

But in so approaching language naturalistically, do we lose something important? Just as Stroud accuses Quine of changing the subject in naturalizing epistemology, one might think that Quine merely changes the subject by naturalizing our theory of language. Although Quine was dismayed that the intensional had become “addictive” for Carnap in his late turn to semantics (TDR, 56), many philosophers view Quine’s own resolute extensionalism to be an unappealing (and ultimately unmotivated) extremism. Might rejecting meanings from our theory be too difficult a pill to swallow? We are typically confident, for example, that we know exactly what our close friends mean by a pointed phrase. Is the pared down naturalist ersatz that we are just better at predicting friends’ speech behavior adequate to capture our lived experience as language users? Faced with the choice between either excising skepticism about communication with Quine or accepting Carnap’s explication that humans strive to know each other’s thoughts

---

<sup>62</sup> Quine compares language users to topiary bushes (WO, 8). Despite being composed of a unique array of branches (i.e., despite inquirers having unique cognitive makeups and so intricately differing conceptual schemes) each has an identical shape (i.e., there is a parity between their dispositions to responsively utter, which constitutes the ability to converse).

while always lacking certain knowledge that they understand each other, Carnap's view may still tempt.

One way of turning this lingering dissatisfaction into an argument is to claim that the naturalist approach to language is *incapable* of capturing some feature of our linguistic discourse, as Gary Ebbs has recently done. In the next section, I will develop a response on Quine's behalf to Ebbs' objection, along these lines, that naturalism cannot account for our dispute-resolving practices.

#### **4. Gary Ebbs and Skepticism About The Use of Our Words**

There is no fact of the matter of our interpreting any man's ontology in one way or...in another. Any man's, that is to say, except ourselves. (TTPT, 23)

In *Rule-Following and Realism*, Ebbs aims "to loosen the hold of [naturalism]" by "develop[ing] a different way of looking at our linguistic practices" (1997, 1). From our perspective as *participants* using language, Ebbs claims, we see that naturalism cannot account for our dispute-resolving practices, because it mischaracterizes the norms that govern our linguistic behavior. Since we can "use [the participant perspective] to criticize and reject...Quine's skepticism about meaning" (92), Ebbs characterizes Quine's naturalism as "mistaken" (4).<sup>63</sup>

---

<sup>63</sup> I admire Ebbs' book, particularly his recognition that philosophical disputes about logic, objectivity and intelligibility are shaped by metaphilosophical, methodological disputes. But I believe him wrong to say that "[w]orries about philosophical relativism evaporate...when we see that some ways of looking at things are better than others, and that we can do no better than to occupy some standpoint or other, the best one we can articulate at the time" (90). Ebbs seeks to replace naturalism with a "better" *participant* picture—"better" because it does not "prevent us from properly describing our linguistic practices" (65). But in my view, no single way of looking at things is "best," or "better" for *all* of our philosophical inquiries into language. So I do not think that Quine's naturalism is "mistaken," but rather that some mistakenly take it to be the sole picture available. I should like to say that worries about philosophical relativism evaporate when we see that some ways of looking at *this thing* are better than others, and that we may legitimately choose the best standpoint we can currently articulate for our current purposes. I will return to this point, and how my methodological position contrasts with Ebbs', in chapter four.

Normativity is a divisive issue between naturalists and their opponents. The point that Ebbs wants to press against Quine parallels the one that John McDowell wants to press against Crispin Wright's account of rule-following (1984, 335-336). According to McDowell, Wright's conception of language fails to adequately capture normativity:

[T]here is nothing but verbal behavior and (no doubt) feelings of constraint. Presumably people's dispositions to behavior and associated feelings match in interesting ways; but at this ground-floor level there is no question of shared commitments—everything normative fades out of the picture. (341)

Since Ebbs explicitly engages Quine's texts, I will focus upon him in this section, and argue that his is an instructive misreading of Quine. Quine's naturalist conception of our "home" language can adequately account for the practices that worry Ebbs. Moreover, viewing our disputes naturalistically excises a third avenue of skeptical doubt. By exposing Ebbs' mistake, we can reconstruct Quine's account of normativity and rational inquiry, despite his rejection of Frege's constitutive view of the relationship between logic and cognition. This work suggests that Wright could similarly rebuff McDowell.

Ebbs interprets Quine's conception of language as fundamentally "idiolectical." We saw in section three that naturalism reduces "communication" between inquirers to their use of a system of noises to pursue shared projects. Inquirers who have the ability to communicate with each other approximately share their linguistic dispositions. The approximation precludes a precise definition of when two speakers "share" a language. "We must keep in mind," Ebbs concludes, that "for Quine a speaker's 'language' consists in his 'present dispositions to verbal behavior' (WO, 27)" (1997, 55). Ebbs proceeds to reconstruct Quine's arguments making this idiolectical conception explicit.

An important example is Ebbs' interpretation of ontological relativity. Quine famously argues that the indeterminacy of translation makes reference inscrutable. We typically translate

each other homophonically, taking others to be referring to the referents of our own expressions. But, workable translation manuals could be designed that use “proxy functions” to permute each of our objects and general terms onto proxies (OR, 55-58). We would still have the capacity to engage in smooth dialogue with others whom we translated as systematically referring to proxies of our own referents.<sup>64</sup> Quine concludes that there is no fact of the matter about another’s ontology; ontology is not absolute, but relative to a manual of translation (53).

However, Quine recognizes that this may render our own capacity to refer dubitable:

we seem to be maneuvering ourselves into the absurd position that there is no difference...between referring to rabbits and referring to rabbit parts or stages...Reference would seem now to become nonsense not just in radical translation but at home. (OR, 47-48)

The worry is that we can apply proxy functions to our own utterances just as easily as to others’. Quine elsewhere notes that we each endorse *disquotational* instances of referential paradigms, sentences where we use the same sentences, adjectives, and names respectively to fill each blank (NTR, 134):

‘ \_\_\_\_\_ ’ is true if and only if \_\_\_\_\_.  
 ‘ \_\_\_\_\_ ’ is true of every \_\_\_\_\_ thing and nothing else.  
 ‘ \_\_\_\_\_ ’ names \_\_\_\_\_ and nothing else.

The obviousness of instances of these paradigms demonstrate their centrality to our web of belief, and clarify how “truth” and “reference” apply in our language. But once we acknowledge that objects in our ontology can be mapped onto proxies specified by a proxy function, then we can generate a set of deviant, *proxy* instances of these paradigms that also preserve all of our speech dispositions.<sup>65</sup> Now we may panic: when I use a name, do I *really* name the object I

---

<sup>64</sup> Suppose our proxy function maps objects in our ontology to space-time regions *n* miles to their west, and we construct proxy predicates P\* for each of our predicates P that are true of exactly the space-time region proxies that are *n* miles to the west of those objects of which our predicates are true. Then when you use the words I use to refer to objects, I may translate you as referring to their space-time proxies, while nevertheless continuing to successfully engage in shared projects with you.

<sup>65</sup> To continue the example from the previous footnote, one instance of the third paradigm will be: “‘rabbit’ names every space-time region *n* miles to the west of a rabbit and nothing else.”

think I do, or only its proxy? I cannot be naming both, since objects are not their proxies, so a single name cannot name *both* a given object *and* its proxy *and* nothing else. But there seems to be nothing in virtue of which I count as referring to an object rather than its proxy. My *belief* that I endorse disquotational instances of the referential paradigms cannot suffice, because if my words do not refer as I think they do, then what seems like a disquotational instance to me may actually refer to a proxy. But on the other hand, it is surely absurd that I cannot distinguish objects from proxies, because I must be referring to them separately when talking about them.

Quine suggests we overcome this panic “in practice” by “acquiescing in our mother tongue and taking its words at face value” (OR, 49); i.e., we take our words to have their disquotational referents. But he does not intend this as pragmatic advice for resisting a genuine semantic problem. Rather, he believes that there *is* no problem once we view matters naturalistically. Naturalists acknowledge that *our own* ontology is not intelligible absolutely, but only against the backdrop of our language and our ongoing theory. We may “acquiesce” about the disquotational reference of terms in our language because we *use* that language (and its tacit disquotational referents) to construct both the proxy and disquotational instances of the referential paradigms in which we merely *mention* the expressions of our language. Our practice is prior to the semantic theory we develop for our practice; use is prior to mention. So, despite the inscrutability of reference, speaking from within our language, we can use our words to successfully make assertions about the things to which we are referring. And there is no position of “cosmic exile” beyond our language, or outside of our theory, which we can occupy to question to what our words absolutely refer (WO, 275).

Ebbs interprets Quine here to be making a claim about each individual’s idiolect. Each of us must tacitly choose a disquotational translation manual for our own idiolect before we “can



*use* our words to state a reference scheme for our own idiolect” (1997, 58). So Ebbs thinks that the shared “mother tongue” in which Quine says inquirers acquiesce is ultimately rhetorical flourish. Strictly speaking, Ebbs’ Quine holds that each inquirer acquiesces in her own private idiolect to questions about ontology. Furthermore, it is only from within her private idiolect that each inquirer can make assertions.

Ebbs thinks that this idiolectical conception of assertion cannot account for our dispute-resolving practices. He complains that “in the glare created by Quine’s behavioristic picture of language use all ordinary talk of meaning looks flimsy and insecure” (62). But Ebbs insists, *contra* Quine, that philosophers of language are not scientists observing a mysterious alien practice. They participate in the very practice they seek to elucidate. Once we adopt the participant perspective, Ebbs thinks that “there is no good reason to accept Quine’s...thesis that between any two speakers there are countless radically different ‘translations’ that capture all that is objective to meaning” (172). Rather, we see that the shared norms governing our practice are the objective ground of our talk of “meaning.”

To defend this claim, Ebbs compares discussions of our practice of using language to discussions of our practice of dancing.<sup>66</sup> Suppose that some people discuss a Scottish dance, the eightsome reel, over dinner. They each learnt how to dance it as children, but now cannot remember one of its steps. They resolve to dance the reel together after pudding. In one case, the diners reach the forgotten step and smoothly continue dancing. They can retroactively describe their activity and each feel confident that *this* description describes how the eightsome reel is danced. In another case, the diners fall out of step with each other at the key moment. Perhaps some of the diners now try to persuade the others that *they* know how the dance is correctly danced. If the group is persuaded, then this case proceeds as the last. In both cases,

---

<sup>66</sup> This analogy is originally R.M. Hare’s (1992).

since “the eightsome reel” had a meaning for the diners prior to their dance, their dancing counts as *clarifying* its meaning and their description is a normative characterization of *how the eightsome reel is danced*.<sup>67</sup>

      Ebbs takes the point of his story to be that “if we are to understand the dispute and its resolution, we cannot restrict ourselves to descriptions available to someone who is unable to recognize or participate in performances of the eightsome reel” (249). For if a group of anthropologists unfamiliar with the eightsome reel were to witness the dancers’ dispute and describe their post-dinner activity and resolution, they would only be able to conclude that the dancers *have agreed to call this activity “the eightsome reel.”* Since there is no meaning to “the eightsome reel” for the anthropologists prior to their observation of the dancers, their characterization of the dance is limited to a non-normative description of what they have observed.

      Ebbs compares having an idiolectal conception of language to approaching the dancing diners like an anthropologist. In his view, naturalists cannot give normative characterizations of how language *is to be* used, but only descriptive characterizations of how those whom they observe *are disposed to* use language. These descriptive characterizations are really phrased in each naturalist’s unique idiolect, and so, constitute part of an *individual’s* linguistic theory of other speakers. If two anthropologists arrive at different descriptions of the dancers, each is entitled to hold her own description correct, so long as it fits the behavior *she* has observed. Similarly, if two naturalists arrive at different descriptions of how language is used, they may each hold their own theory correct. Just as the anthropologists “miss an important aspect of the

---

<sup>67</sup> The remaining case of *unresolved* dispute, where the group is not persuaded about how to dance the dance, does not concern us here.

practices they observe” (308)—namely, what activity the eightsome reel *is*—Ebbs charges the naturalist with failing to recognize that

the agreements and disagreements that we express in our everyday and scientific inquiries are rooted in our *shared* practical ability to make archetypical applications of our words, and in our commitment to resolve disputes about the conditions for properly applying our words by appealing in part to that shared practical ability. (309, my emphasis.)

Ebbs concludes that philosophers of language must occupy a participant perspective if they are to properly account for our practice of resolving disputes together. We do not create *individual* theories of other’s acceptable language use. The labor of acceptable language use is rather divided between all speakers (308). We hold ourselves accountable to each other, not merely to the data that we have each collected.

But this objection will not convince the naturalist. Ebbs tells us that the diners learnt the eightsome reel as children. But the naturalist should ask what justifies the diners’ tacit assumption that there was a *single* dance that they learned, which they have now correctly “clarified” and described. How do the diners know that their instructors did not teach them different dances? How do they know that the eightsome reel does not have variations?

Ebbs will reply that the normative characterization at which the dancers arrive is fallible. He will admit that they may be mistaken about the eightsome reel, despite their current agreement. But the naturalist will now ask what “correctness” for the dancers amounts to if their current description can be “mistaken.” Although the dancers can undertake further investigation to increase their *confidence* in their assumption that there is a single dance which they are characterizing (perhaps by talking to their old teachers or consulting independent sources which describe Scottish reels), they can *never* be *certain*. And, once again, a skeptic can exploit this epistemic gap to doubt the diners’ putative normative characterization of the dance. However confident the dancers feel that *this* is the eightsome reel, they do not *know* it is—nor even that

there *is* a single dance to be known. Stepping back from the analogy, so long as philosophers of language occupy the participant perspective and assume that each word of a language *possesses a meaning*, and that through discussion with other competent language users we can arrive at *fallible* normative characterizations of these meanings, there is room to skeptically doubt that speakers know what their words mean.

By denying that our words have meanings to which our descriptions of language use are accountable, Quine excises these skeptical doubts. He thinks our theoretical descriptions of language use are only accountable to the scientific standards to which all of our theories are accountable. But it is a mistake to think that Quine thereby eliminates the linguistic obligations we incur when we reflect on acceptable language use in our own speech community.

To see why, consider again the anthropologist's record of the post-dinner festivities. Having noted the diners' agreement to call their activity "the eightsome reel," she erects a standard that generates obligations for future speech behavior. For example, if one of the dancers (or another anthropologist) later calls some new dance "the eightsome reel," she may exhibit her record and demand a justification for the deviant usage. The obligations about how to use the phrase "the eightsome reel" that are generated by her record are defeasible, but not *wantonly* so. They may be legitimately challenged by alternative hypotheses that meet the scientific standards countenanced by naturalist inquirers.<sup>68</sup> But an interlocutor who merely courts eccentricity can be justly dismissed.

Similarly, the theories that naturalists develop about the appropriate usage of particular words (defeasibly) oblige all inquirers. Any challenges to these linguistic theories (which are

---

<sup>68</sup> For example, the anthropologist's interlocutor may object that since she has friends who also claim to know the dance but dance it quite differently, the sample upon which the anthropologist constructed her theory is too small.

largely uncontroversial) will be evaluated at the scientific tribunal. So Ebbs is wrong to think that the naturalist lacks the means to account for our practice of resolving disputes together.

But although Ebbs' objection will not convince the committed naturalist that she is mistaken, might it convince the philosopher of language to refuse naturalism?<sup>69</sup> I have been arguing that Quine's naturalism has the power to excise skeptical challenges about our knowledge of the external world, of successful communication, and of what our words mean. But if it does so by carelessly wielding a machete through our core intuitions about what we want our theories of knowledge and language to explain, perhaps Ebbs' reader will join him in believing that "properly describing" our linguistic practices requires embracing the participant perspective, even if doing so makes those skeptical threats intelligible (65).

However, I began this chapter by claiming that naturalism is rightly viewed as a paring knife, not a machete, a metaphor intended to convey the ways naturalists' critics often mistake what is and what is not eliminated by eliminating "meanings" from our theory. Ebbs, for example, finds the naturalist account of our dispute-resolving practices wanting. But he only does so because he mistakenly thinks that "idiolects" bear the weight of Quine's theory. In fact, Quine has a more nuanced conception of language.

Ebbs is right that Quine believes the relation of "sharing a language" only informally applicable, since no two speakers are familiar with exactly the same vocabulary, and nor are they disposed to use words in their vocabularies in just the same circumstances. But this does not entail that each speaker's capacity for using language reduces to her present dispositions to verbal behavior. This idiolectal description of a speaker misses Quine's recognition that

---

<sup>69</sup> Ebbs admits in the final section of his book that "[n]o one who has ever been in the grip of Quine's picture of meaning and assertion will be immediately convinced by the foregoing reasons for resisting it. Our philosophical commitments and temptations, especially if they are deeply rooted and systematic, have a momentum of their own" (309). So perhaps, despite his earlier declaration that he will show naturalism "mistaken," his real goal is to convince non-naturalists to remain non-naturalists.

language is “our social art,” a fundamentally shared activity that is identifiable in a community of speakers *and* listeners whose activities are coordinated. The idiolectal description of a speaker’s language that Ebbs extracts from Quine is just an idealization that Quine employs to simplify the presentation of his theory of language. But Quine knows that there is no private language (OR, 27). His idiolectal idealization is embedded in his intersubjective naturalist account of language acquisition.

In Quine’s view, with the guidance of parents and teachers who reward and rebuke our burgeoning efforts as needed, we acquire our *home language* socially as children. In so doing we add strands to our webs of belief about how the words and sentences we begin to utter apply to things and facts in the world. Since this education admits of considerable variation, any two children will acquire slightly different linguistic competencies, but within geographic and cultural boundaries are able to use their verbal dispositions to coordinate their activities, to “communicate” with each other, and to count as sharing a language: “we demarcate our practical speech community, for particular given purposes, as the community in which all dialogue that is concerned with those purposes runs smoothly and effectively (R, 311).<sup>70</sup> When Quine describes “acquiescence in our mother tongue” as the way to resolve the appearance of paradox in ontological relativity, he refers to our *shared* home language, not, as Ebbs claims, our individual idiolects. Our individual utterances are meaningful in our community because we share our home language.<sup>71</sup> This casts Quine’s account of assertion and our dispute-resolving practices in a different and more attractive light.

---

<sup>70</sup> We may later expand our linguistic competency by learning how to translate foreign languages, which Quine terms “alien” languages. Quine understands the learning of a foreign language on his translational model.

<sup>71</sup> Thus as I read Quine, our shared home language becomes a pared down, extensionalist, but nonetheless social correlate to Carnap’s inflated, intensional assumption that speakers share a fixed interpretation of a metalanguage, which deprives the skeptic of the vocabulary needed to press radical objections.

The conclusion of Quine's reflections on ontological relativity is that there is no perspective external to the *home language* from which inquirers can question their ontology.<sup>72</sup> Within the home language, using its terms, we can study, evaluate, and continually refine the ontological commitments of our theories. Since our assertions are made from within this home language, we are accountable to other speakers for the meaningfulness of our utterances.

The idiolectal description of our unique speech dispositions is a theoretical posit phrased *using* the home language when we decide to investigate our ability to converse. But since this theoretical grasp of our unique idiolect depends upon our shared home language, it cannot be occupied to impugn it. We employ our home language both when considering its reference and when constructing alternative possibilities about deviant reference. So Quine thinks we only grasp *factuality* from a perspective within our home language: "factuality, like gravitation and electric charge, is internal to our theory of nature" (TTPT 23).<sup>73</sup> It follows that we must ultimately acquiesce in our home language to the disputes about ontology that we phrase *in* that language, *not* our unique idiolects.<sup>74</sup>

Ebbs was worried, justifiably, that, under the idiolectal conception of language, linguistic disputes become peculiarly individualized. If each individual is only able to assert from within her own idiolect, then any differences between her theory of acceptable language

---

<sup>72</sup> In a recent paper, Thomas Ricketts develops a persuasive interpretation of ontological relativity which emphasizes the home language (2010). Ricketts notes firstly that the disquotational instances of the referential paradigms are privileged because one learns them in learning the home language, as opposed to Ebbs' claim that each speaker endorses instances of them for her own idiolect. Furthermore, Ricketts argues that in any conversation where one uses a non-homophonic translation manual to translate another speaker, one must "respond directly" to the home language sentence that one understands which the non-homophonic translation replaces for the interlocutor's sentence, concluding that "in these two ways...reliance on the...homophonic translation manual is inescapable for speakers of the home language" (14). (Ricketts here credits Peter Hylton [2007].)

<sup>73</sup> It is problematic to interpret Quine's use of "our" in this quote—as elsewhere in his writings. In my view, Quine here means *our shared* theories of nature, not our individual theories of nature.

<sup>74</sup> C.f. Ricketts' argument that Quine understands the disquotational instances of the referential paradigms to give a "formal mode specification" of "facts of the matter": "the conclusion here is *not* that, in contrast to alien languages, there is a fact of the matter about reference in the home language. Rather, because of the use of a disquotationally explicated reference-predicate in characterizing 'facts of the matter,' this notion of *facts of the matter* is not applicable to home language reference" (2010, 21).

use and others' can only generate optional changes she may make to that theory if she wishes to improve her future efforts at communication. Others cannot show her linguistic theory *incorrect*, but only *unfruitful* if used in future conversations with them. But for Quine, each individual is only able to assert on the basis of having learned her home language with others. Although no two speakers are identically disposed to use exactly the same expressions within the home language, there is no part of an individual's idiolect that is unshared with some other speaker, for it is only *sharing* linguistic dispositions that make our words meaningful. So, in Quine's view, disputes about how home language words are to be used are not individualized. If an individual cannot justify her linguistic theory in the face of challenges, it is to be judged incorrect according to the shared scientific standards that guide our inquiry.

I concluded section one with a number of questions facing Quine's naturalist conception of the relationship between logic and cognition, given his rejection of Frege's constitutive view. Firstly, does admitting that the laws of canonical logic can alter under pressure from our theories force Quine into sociologism about logic, where objective truth changes with our evolving patterns of reasoning? Secondly, what norms can we know our interlocutors follow, and how do we know that we are agreeing or disagreeing rather than talking past one another? Thirdly, what conception does the naturalist have of normativity, and is it sufficient to model inquiry? Having improved our grasp of Quine's naturalism, we are now positioned to answer these questions.

Quine's understanding of the home language allows him to avoid sociologism about logic. Whereas we determine which logic to canonize by observing and refining the inferential patterns upon which our discourse depends, the application of truth and reference is grounded not on observation of each other but rather in the obviousness of instances of the disquotational schemata to anyone who learns the home language. So "truth" is not defined as agreement



according to the logical inference relation our inquiring community currently adopts (and nor is “reality” defined as our current theory of reality). Although the word “truth” is part of our inferential practice—only (fallibly) judged by inquirers who engage in our practice—Quine thinks it “transcends” our theory *semantically*, because the truth predicate for a language cannot be defined within that language, and *doctrinally*, because “science in even the broadest sense” of “informed belief” will not take a stance on the truth or falsity of every sentence within a language (WWD, 78). Comprehensive knowledge of all truths is an unrealizable *goal* for our logically certified scientific inquiry, not its *product*. So truth is “[a]n ideal of pure reason, yes, and hallowed be its name” (78).<sup>75</sup>

Secondly, we saw in section three that Quine’s behaviorist ersatz of “communication” robbed any radical skeptical doubt that we might be merely talking past one another instead of agreeing or disagreeing. Earlier, I noted that naturalized epistemology embraces the mutual containment of epistemology and science. We find another application of this in the naturalist theory of language acquisition and use. Our scientific theory of evolution explains how we acquired the capacity to use language together, and our ability to use language together is the basis of our ongoing scientific theories. Quine argues that natural selection accounts for our “meeting of minds” about how to carve up our (unshared) sensory experiences, and so, also for our ability to acquire a shared home language (IYI, 486). Each human organism has the capacity

---

<sup>75</sup> This jolly remark might seem surprising given Quine’s response to skepticism about the external world. For if truth is doctrinally transcendent, surely reality must be too. If, like Lars Bergström, we *identify* the truths of reality with the truths of one’s current theory, the skeptic is denied concepts of “truth” or “reality” external to her theory with which to coherently present her challenge (1994, 430). But, in accepting that an individual’s concepts of “truth” and “reality” name goals which transcend her current theory, has Quine reopened the door to Stroud’s skeptic? The answer is no, because Quine can maintain that, as we are naming goals from within our theory, these goals are only evaluated as part of our theory. He will naturalistically rephrase the skeptic’s doubt that our current theory even *approximates* the “reality” which is its aim as wondering whether current science, despite its simplicity and impressive predictive power, is badly wrong. Quine will agree that future experiences may show that part of our theory *is* badly wrong. Yet, once again, to doubt scientific realism on this basis alone—on science’s fallible grasp of its goal—is “overreacting” (RS, 475).

to determine perceptual similarities in its environment (PTF, 475), and the human species has collectively evolved to be in “intersubjective harmony,” to find similar aspects of the world salient. Over time, natural selection has ensured a continued parallelism between the parts of the world we find perceptually similar. This parallelism entails that fellow home language speakers will balk at a deviant application of a home language sentence to the world. The naturalist inquirer recognizes that she is fallible about the application of words and that she may not have the best explanatory theory. She demonstrates her sensitivity to this epistemic situation by continually refining (and accepting others’ refinements of) her theories. In this way, our theory of evolution accounts for the shared principles of inference that underwrite human inquiry.

I believe that Quine can likewise appeal to the theory of evolution to resist a version of the objection that I earlier described McDowell raising against Wright. Even having acknowledged Quine’s emphasis on the home language rather than individual idiolects, one might still think his conception of language similarly involves “nothing but verbal behavior and...feelings of constraint” among members of a community. But for Quine, that is certainly not to say that there is “no question of shared commitments,” because, as our theory of evolution tells us, our parallel dispositions are not merely “interesting” but are a *condition* for us to have the capacity to use language at all. It is our shared similarity standards of stimuli in the environment that allow us to talk together about the world and to commit to developing theories of its nature together.

These reflections point the way to answering the final question. The reason that Quine says little about logical or epistemic norms is that he thinks that what philosophers have called “normativity” is elaborated in the substance of scientific theories. Logical truths, for instance, oblige inquirers in exactly the sense that any scientific truth obliges inquirers; since they are true,

truth-seeking inquirers ought to judge them true. What we might be tempted to call logic's unique "normative force" is really just a function of logic's generality across scientific theories. Similarly, "epistemic norms" are really just a way to characterize in general terms, from within our ongoing theory, the types of chains of reasoning we endorse. We lose nothing by naturalistically expressing such "norms" in descriptive terms by elaborating our current methodology.

In particular, each individual inquirer's openness to revision grants the naturalist a conception of objectivity and allows her to explain our dispute-resolving practices. In the context of our theory of evolution, we see that the goal of truth is shared by all speakers of the home language, who—as we do—come to find the disquotational instances of the truth and reference paradigms obvious in learning it. Our aim as inquirers is hence not private and subjective, but shared and objective. In cases of theoretical disputes, including disputes about correct language use, the onus rests upon the individual who judges against the group to make her case before the scientific tribunal. And in this context logic comes to the foreground, not as our attempt to uncover the constitutive prerequisites of thought to which we are all tacitly committed, but as our attempt to refine our understanding of the principles that we use to evaluate our theories. It is in our interest as inquirers to choose a translation of others under which they reason according to our canonical logic (and so, under which they acknowledge their obligation to its "norms"), because doing so promotes our mutual evaluation of each other's—and improvements of our own—theories.

In conclusion, I have argued that Quine's naturalism is far more subtle than is usually acknowledged. It pares down traditional epistemological discussions by removing the vocabulary in which skeptical threats about the external world, successful communication, and

the use of our own words can be framed. Naturalism also defines a clear philosophical program for logic. It casts logic as central for clarifying our theories, as clarifying our reasons for adopting theories, and for making explicit our standards for testing disputed claims. Quine adopts Frege's insight that codifying the principles of inference is a matter of codifying the rules of our demonstrative practice, but rejects his constitutive view and constructs a holistic theory of justified beliefs. Any one of our beliefs can be scrutinized and potentially found lacking, but one theory is not better than another according to an epistemically inaccessible, unquestionable, and unrevisable standard. Naturalized philosophers recognize there are no such standards. In Quine's view, rather than chasing such ghosts, the philosophical community ought to join scientists in refining the only standards we have to go on—our fallible, but nevertheless currently effective, scientific standards.

## **Bibliography**

- Awodey, Steve and Carstein Klein (eds). *Carnap Brought Home: The View from Jena*. Chicago and La Salle: Open Court, 2004.
- Barrett, Robert and Roger F. Gibson, eds. *Perspectives on Quine*. Oxford: Blackwell, 1990.
- Bell, David and Willhelm Vossenkuhl, eds. *Science and Subjectivity: The Vienna Circle and Twentieth Century Philosophy*. Berlin: Akademie Verlag, 1992.
- Berger, Alan. "A Central Problem for a Speech-Dispositional Account of Logic and Language." Barrett and Gibson 1990, 17-35.
- Bergmann, Gustav, *The Metaphysics of Logical Positivism*. London: Longman's Green & Co., 1954.
- Bergström, Lars. "Quine's Truth." *Inquiry*, 37, 4 (1994): 421-435.
- Beth, E.W. "Carnap on Constructed Systems." Schilpp 1963, 469-502.
- Birkhoff, Garrett and John von Neumann. "The Logic of Quantum Mechanics." *Annals of Mathematics*, 37, (1936): 823-843.
- Carnap, Rudolf. "E. W. Beth on Constructed Languages." (BCS) Schilpp 1963, 927-933.  
\_\_\_\_\_. *Foundations of Logic and Mathematics*. (FLM) Chicago: University of Chicago Press, 1939.  
\_\_\_\_\_. "Intellectual Autobiography." (IA) Schilpp 1963, 3-84.  
\_\_\_\_\_. *The Logical Foundations of Probability*. (LFP) 2<sup>nd</sup> ed. Chicago: University of Chicago Press, 1967.  
\_\_\_\_\_. *The Logical Syntax of Language*. (LSL) 1937. Trans. Amethe Smeaton. London: Kegan Paul, Trench, Trubner & Co., trans. of *logische Syntax der Sprache*. Vienna: Julius Springer Press 1934. Illinois: Open Court Press, 2002.  
\_\_\_\_\_. "My Conception of Semantics." (MCS) Schilpp 1963, 900-905.  
\_\_\_\_\_. *Meaning and Necessity*. (MN) Enlarged ed. Chicago: University of Chicago Press, 1958.  
\_\_\_\_\_. *Philosophy and Logical Syntax*. (PLS) London: Kegan Paul, Trench, Trubner & Co, 1935.  
\_\_\_\_\_. "Quine on Logical Truth." (QLT) Schilpp 1963, 915-922.  
\_\_\_\_\_. "Strawson on Linguistic Naturalism." (SLM) Schilpp 1963, 933-940.  
\_\_\_\_\_. "The Task of the Logic of Science." (TLS) Trans. of *Die Aufgabe der Wissenschaftslogik*. Vienna: Gerold, 1934. *Unified Science*. ed. Brian McGuinness, Dordrecht: D. Reidel Publishing Co., 1987, 46-66.  
\_\_\_\_\_. "Testability and Meaning." (TM) *Philosophy of Science*, Vol 3., No. 4, 1936; Vol 4., No. 1, 1937.

- Creath, Richard. "Quine on the Intelligibility and Relevance of Analyticity." Gibson Jr. 2004, 47-64.
- Davidson, Donald and Jaakko Hintikka, eds. *Words and Objections*. Rev ed. Dordrecht: D. Reidel Publishing Co., 1975.
- Dreben, Burton. "Putnam, Quine, and the Facts." *Philosophical Topics*, 20, 1 (1992): 293-315.  
 \_\_\_\_\_. "In Mediis Rebus." *Inquiry*, 37, 4 (1994): 441-447.  
 \_\_\_\_\_. "Quine on Quine." Gibson Jr. 2004, 287-293.
- Ebbs, Gary. *Rule-Following and Realism*. Harvard: Harvard UP, 2001.
- Field, Hartry. *Saving Truth from Paradox*. Oxford: Oxford UP, 2008.
- Frank, Philipp. "The Pragmatic Components in Carnap's 'Elimination of Metaphysics.'" Schilpp 1963, 159-164.
- Friedman, Michael. *Reconsidering Logical Positivism*. Cambridge: Cambridge UP, 1999.
- Gibson Jr., Roger F., ed. *The Cambridge Companion to Quine*. Cambridge: Cambridge UP, 2004.
- Glock, Hans. *Quine and Davidson on Language, Thought and Reality*. Cambridge: Cambridge UP, 2003.
- Goldfarb, Warren and Thomas Ricketts. "Carnap and the Philosophy of Mathematics." Bell and Vossenkuhl 1992, 61-78.
- Guttenplan, Samuel, ed. *Mind and Language*. Oxford: Clarendon Press, 1975.
- Hahn, Lewis, ed. *The Philosophy of Donald Davidson*. Chicago: Open Court, 1999.  
 \_\_\_\_\_ and Paul Schilpp, eds. *The Philosophy of W.V. Quine*. Exp. ed. Illinois: Open Court Press, 1998.
- Hanna, Joseph F. "An Explication of 'Explication.'" *Philosophy of Science*, 35, 1 (1968): 28-44.
- Hare, R. M. "Philosophical Discoveries." *Mind*, 49, (1960): 145-162. Rpt. in Rorty 1992, 206-217.
- Hilgevoord, Jan and Jos Uffink. "The Uncertainty Principle." *The Stanford Encyclopedia of Philosophy* (Spring 2011 edition), Edward N. Zalta (ed.), forthcoming URL = <http://plato.stanford.edu/archives/spr2011/entries/qt-uncertainty/>.
- Hylton, Peter. *Quine*. New York: Routledge Press, 2007.
- Isaacson, Daniel. "Carnap, Quine and Logical Truth." Bell and Vossenkuhl 1992, 100-130.

- Kuratowski, Kazimierz. "Sur la notion de l'ordre dans la théorie des ensembles." *Fundamenta Mathematicae*, 2 (1921): 161-171.
- Lévy-Bruhl, Lucien. *How Natives Think*. Trans. Lilian Clare. Trans. of *Les fonctions mentales dans les sociétés inférieures*. Paris: Felix Alcan, 1910. London: George Allen and Unwin, Ltd., 1926.
- McDowell, John. "Wittgenstein on Following a Rule." *Synthese*, 58 (1984): 325-363.
- Margolis, Joseph. *The Unraveling of Scientism: American Philosophy at the End of the Twentieth Century*. Cornell: Cornell University Press, 2003.
- Putnam, Hilary. "How to think Quantum-logically." *Synthese*, 29, (1974): 55-61.  
 \_\_\_\_\_. "Meaning Holism." Hahn and Schilpp 1998, 405-426.
- Quine, W. V. "Carnap and Logical Truth." (CLT) Schilpp 1963, 385-406.  
 \_\_\_\_\_. *Confessions of a Confirmed Extensionalist and Other Essays*. Eds. Dagfinn Føllesdal and Douglas Quine. Cambridge: Harvard UP, 2008a.  
 \_\_\_\_\_. "Confessions of a Confirmed Extensionalist." (CCE) Floyd and Shieh 2001, 215-222. Rpt. Quine 2008a, 498-506.  
 \_\_\_\_\_. "Epistemology Naturalized." (EN) Quine 1969, 69-90.  
 \_\_\_\_\_. "Four Hot Questions in Philosophy." (FHQP) *New York Review of Books*, Feb 14, 1985, 13. Rpt. in Quine 2008b, 206-210.  
 \_\_\_\_\_. *From a Logical Point of View* 2<sup>nd</sup> Rev. ed., Cambridge: Harvard UP, 1980.  
 \_\_\_\_\_. "I, You and It: An Epistemological Triangle." (IYI) *Knowledge, Language and Logic: Questions for Quine*. Eds. Alex Orenstein and Petr Kotatko. Boston: Kluwer Academic Publishers 2000, 1-6. Rpt. in Quine 2008a, 485-492.  
 \_\_\_\_\_. "Natural Kinds." (NK) Quine 1969, 114-138.  
 \_\_\_\_\_. "The Nature of Natural Knowledge" (NNK) Guttenplan 1975, 67-81. Rpt. Quine 2008a, 257-271.  
 \_\_\_\_\_. "Notes on the Theory of Reference" (NTR) in Quine 1980, 130-138.  
 \_\_\_\_\_. "Ontological Relativity." (OR) Quine 1969, 26-68.  
 \_\_\_\_\_. *Ontological Relativity and Other Essays*. New York: Columbia UP, 1969.  
 \_\_\_\_\_. "On What There Is." (OWTI) *Review of Metaphysics*, 2 (1948): 21-38. Rpt. Quine 1980, 1-19.  
 \_\_\_\_\_. *Philosophy of Logic*. (PL) 2<sup>nd</sup> ed. Cambridge: Harvard UP, 1986.  
 \_\_\_\_\_. "Progress on Two Fronts." (PTF) *Journal of Philosophy*, 93, 4, (1996): 159-163. Rpt in Quine 2008a, 473-477.  
 \_\_\_\_\_. *Pursuit of Truth*. (PT) Rev ed. Cambridge: Harvard UP, 1992.  
 \_\_\_\_\_. *Quiddities*. (Q) Cambridge: Harvard UP, 1987.  
 \_\_\_\_\_. *Quine in Dialogue*. Eds. Dagfinn Føllesdal and Douglas Quine. Cambridge: Harvard UP, 2008b.  
 \_\_\_\_\_. "Relativism and Absolutism." (RA) *The Monist* 67, 3 (1984): 293-296. Rpt in Quine 2008a, 319-322.  
 \_\_\_\_\_. "Replies." (R) Davidson and Hintikka 1975. 292-352.

- \_\_\_\_\_. "Responses." (R2) Quine 1981. 173-186.
- \_\_\_\_\_. "Reply to Herbert G. Bohnert." (RHB) Hahn and Schilpp 1998, 93-95.
- \_\_\_\_\_. "Reply to Hilary Putnam." (RHP) Hahn and Schilpp 1998, 427-432.
- \_\_\_\_\_. "Reply to John Woods." (R JW) Hahn and Schilpp 1998, 726-728.
- \_\_\_\_\_. "Reply to Paul A. Roth." (RPR) Hahn and Schilpp 1998, 459-460.
- \_\_\_\_\_. "Reply to Roger F. Gibson, Jr." (RG) Hahn and Schilpp 1998, 684-685.
- \_\_\_\_\_. "Reply to Stroud." (RS) *Midwest Studies in Philosophy VI: The Foundations of Analytic Philosophy*. Eds. P. French, T. Uehling and H. Wettstein. Indianapolis: University of Notre Dame Press, 1981. 473-475.
- \_\_\_\_\_. *The Roots of Reference*. (RR) La Salle, Illinois: Open Court Press, 1974.
- \_\_\_\_\_. "The Scope and Language of Science." (SLS) *The Unity of Knowledge*, ed. Lewis Leary. New York: Doubleday, 1955. Rpt in Quine 1966, 215-232.
- \_\_\_\_\_. "Structure and Nature." (SN) *The Journal of Philosophy*, 89, 1, (1992): 1-9. Rpt. in Quine 2008a, 401-406.
- \_\_\_\_\_. "Speaking of Objects." (SO) *Proceedings and Addressees of the American Philosophical Association*, 31, (1958): 5-22. Rpt in Quine 1969, 1-25.
- \_\_\_\_\_. "Truth by Convention." (TC) *Philosophical Essays for A.N. Whitehead*. ed. O.H. Lee. New York: Longmans, 1936. Rpt. in Quine 1966, 70-99.
- \_\_\_\_\_. *Theories and Things*. Cambridge: Belknap Press, 1981.
- \_\_\_\_\_. "Things and Their Place in Theories." (TTPT) Quine 1981. 1-23.
- \_\_\_\_\_. "Three Indeterminacies." (TI) Barrett and Gibson 1990. 1-16.
- \_\_\_\_\_. "Two Dogmas of Empiricism." (TDE) *Philosophical Review*, 60, 1 (1951): 20-43. Rpt. in Quine 1980, 20-46.
- \_\_\_\_\_. "Two Dogmas in Retrospect." (TDR) *The Canadian Journal of Philosophy*, 21 (1991): 265-274. Rpt. in Quine 2004, 54-63.
- \_\_\_\_\_. *The Ways of Paradox and other essays*. New York: Random House, 1966.
- \_\_\_\_\_. "Where Do We Disagree?" (WWD) Hahn 1999. 73-79.
- \_\_\_\_\_. *Word and Object*. (WO) Cambridge: MIT Press, 1960.
- \_\_\_\_\_ and Donald Davidson. "Exchange Between Donald Davidson and W.V. Quine Following Davidson's Lecture." (EDQD) *Theoria* 60, (1994): 226-231. Rpt Quine 2008b, 152-156.
- \_\_\_\_\_ and Joseph Ullian. *The Web of Belief*. (WB) Rev. ed. New York: Random House, 1978.
- Ricketts, Thomas. "Frege, Carnap and Quine: Continuities and Discontinuities." Awodey and Klein 2004, 181-202.
- \_\_\_\_\_. "Roots of Ontological Relativity." Forthcoming in *American Philosophical Quarterly*, 2010.
- Rorty, Richard, ed. *The Linguistic Turn: Essays in Philosophical Method with two retrospective essays*. 2<sup>nd</sup> ed. Chicago: University of Chicago Press, 1992.
- Schilpp, Paul, ed. *The Philosophy of Rudolf Carnap*. Illinois: Open Court Press, 1963.
- Shapiro, Stewart, ed. *The Limits of Logic*. Aldershot: Dartmouth Press, 1996.



Sorell, Tom. *Scientism: Philosophy and the Infatuation with Science*. London: Routledge, 1991.

Stroud, Barry. "Conventionalism and the Indeterminacy of Translation." Davidson and Hintikka 1975, 82-96.

\_\_\_\_\_. *The Significance of Philosophical Skepticism*. Oxford: Oxford UP, 1984.

Tennant, Neil. "Radical Interpretation, Logic, and Conceptual Schemes." in *Interpretations and Causes: New Perspectives on Donald Davidson's Philosophy*, ed. Mario de Caro. London: Kluwer Academic Press, 1999, 71-94.

Ullian, Joseph. "Quine and Logic." Gibson Jr. 2004, 270-286.

## **Davidson's Humanism: Truth Only Through Interpretation**

Unlike his teacher W.V. Quine, who refined his epistemology by penning a series of books relating his evolving views from the ground up,<sup>1</sup> Ernest Lepore and Kirk Ludwig note that Donald Davidson's philosophy is revealed to us through a range of papers that form a "mosaic" (2005, 5). This makes Davidson peculiarly susceptible to misinterpretation. Taken individually, the conclusions reached in his papers can appear unmotivated or implausible. Nevertheless, Davidson's arguments are buttressed by a conception of philosophical inquiry that is every bit as nuanced as Quine's naturalism. In this chapter, I intend to shed light on what I shall call Davidson's "humanism" by exposing the core commitments grounding his epistemology.<sup>2</sup> In doing so, I will argue that Davidson recovers, and transforms, the privileged relationship to cognition Frege awards, and Quine denies, logic.

Davidson's account of radical interpretation is the focus of section one. I isolate the conception of logic it reveals by bringing his claim that logic is required for interpretation together with his claim that interpretation is required for cognition. I argue that the intersubjectivity of radical interpretation distinguishes this conception of logic from Frege's constitutivism and Quine's naturalism. However, Davidson's arguments for the two claims at the heart of his humanist conception, and for the associated thesis that there are no conceptual schemes, are murky. In section two, I reconstruct Davidson's argument against conceptual schemes to highlight its transcendental structure, and I defend Davidson from Simon Blackburn's challenge that his views presuppose verificationism and the related objection that he

---

<sup>1</sup> Although Quine wrote numerous articles, he sketches his epistemology in *Word and Object* (1960), *The Roots of Reference* (1974), *Pursuit of Truth* (1990), and *From Stimulus to Science* (1995).

<sup>2</sup> Unlike Quine, Davidson never branded his conception of philosophy. Although recent critics have sought to bring Quine and Davidson together under a variety of umbrellas (including naturalism [Björn Ramberg, 2001] and logical pragmatism [Hans Glock, 2003]), I believe doing so only obscures their important differences. I defend my choice to call Davidson's approach "humanism" in section three, and I examine the differences between Quine and Davidson in chapter four.

is unwittingly anthropocentric. In section three, I argue that, on the contrary, Davidson's principled humanism underwrites his distinctive analyses of key epistemic concepts. By privileging the mutual interpretation of human inquirers as not merely our meaning-detecting, but our meaning-*making* practice, Davidson renders grasp of the intersubjective concept of truth a condition for cognition. It is this humanism that motivates this philosopher of language's surprising late claim that, "there is no such thing as a language" (NDE, 265).

### **1. Radical Interpretation and Reading Logic Into Language**

Davidson employs radical interpretation to understand "meaning." His strategy is inspired by Alfred Tarski's for understanding truth. Just as Tarski seeks to illuminate truth by finding the desiderata of a *theory* of truth for a language, so Davidson wants to illuminate meaning by finding the desiderata of a *theory* of meaning for a language (TM, 158; RII, 126). Davidson uses the figure of the radical interpreter to elaborate the additional "empirical constraints" upon a theory of truth for a language in use that would license employing it as a theory of meaning for the language users' utterances (RI, 194).

Tarski argues that a truth theory (or "T-theory") for a language L ought to entail, for each sentence S of L, a sentence that instantiates the T-schema

'S' is T if and only if p

where "'S'" is a structural descriptive name of "S," "p" is a sentence in the language of the truth theory that is synonymous with "S," and "T" becomes the theory's truth predicate for L (1956, 187-188). If the language ML of the truth theory contains the language L, "p" can simply be "S." If ML does not contain L, then "p" must be synonymous with "S." In this way, Tarski relies upon translation—the capacity to capture the meaningfulness of sentences from one

language in the sentences of another—to gain purchase upon truth. Davidson proposes inverting this dependence. Since “to give truth conditions is a way of giving the meaning of a sentence” (TM, 160), he contends that if certain criteria are met, someone who has so constructed a T-theory for L can systematically replace each sentence of her theory with one satisfying the schema

‘S’ means that p

resulting in a theory of meaning for L. In the interesting case where ML does not contain L, the theory will *interpret* L in ML.

Davidson is not proposing an *atomistic* reduction of meaning to truth. He does not think, that is, that any isolated ML-sentence instantiating the form of the T-schema will give both the truth conditions and the meaning of a particular L-sentence. Nor will just any theory entailing an infinite set of ML-sentences possessing this form count as a theory of meaning for L. Tarski showed how to develop theories which coordinated structural components of L-sentences with ML-expressions that translated them. The syntax of any L-sentence could be recursively parsed in ML using the axioms of such theories, entailing an instance of the T-schema with a structural description of the L-sentence on the left-hand side and an ML-translation on the right-hand side. Since any L-sentence could be treated by the theory in this way, inquirers were justified in treating “T” as a truth predicate for L. If we drop Tarski’s requirement that an ML-translation for each L-sentence be given at the outset, then the recursive structure of the T-theory remains. Davidson seeks to exploit this structure to show “how the meaning of a sentence depends on its composition” (SNL, 56). If an ML-speaker knows that her T-theory is a theory of truth for L—that is, if she instead assumes truth from the outset—and if it has finitely many axioms for her to

learn and manipulate, then she can employ those axioms to recursively parse out the meaningfulness of particular L-sentences and interpret L-speakers.<sup>3</sup>

However, it is easy to construct theories that satisfy the syntactic requirements of convention T (such as having a recursive structure), but which codify irrelevant semantic content about L. These deviant T-theories cannot be plausibly used as theories of meaning for L.<sup>4</sup> To block these counterexamples, Davidson argues that one wishing to use a T-theory as a theory of meaning for some language in use must know additional facts about how it was constructed from the available evidence. It is here that radical interpretation comes in. According to Davidson, since the radical interpreter is in a position to judge that her T-theory “can be projected to unobserved and counterfactual cases” (RF, 36), she is entitled to use it as a theory of meaning for the language users she observes.

Davidson models radical interpretation on Quine’s radical translation. Like the radical translator, the radical interpreter is utterly unfamiliar with the language spoken in a native community, and so, must try to break into it solely on the basis of her observations. Yet while the radical translator’s goal is a manual that will allow her and the natives to engage in fluent dialogue, where “fluency” is understood behavioristically as the lack of conversational breakdown between translator and native, the radical interpreter is attempting, more ambitiously,

---

<sup>3</sup> The holism underwriting this conception of interpretation—that an interpreter’s theory does not enumerate isolated sentences of L that she understands, but must be structured so as to allow her to interpret vast swathes of L at once—is the sense of Davidson’s modification of Frege’s context principle that “only in the context of the language does a sentence (and therefore a word) have meaning” (TM, 159).

<sup>4</sup> J. A. Foster argues that deviant T-theories undermine Davidson’s program (1976, 13-14). For example, a T-theory for English might entail the sentence “‘The check is in the post’ is T iff the check is in the post and snow is white.” Since snow is, in fact, white, the biconditional will return the correct truth conditions for the contingent English sentence “the check is in the post,” and hence, the theory entailing this sentence might pass as a T-theory and the “T” it defines might be a legitimate truth predicate for English. But this sentence plainly fails to give the *meaning* of “the check is in the post.” Foster concludes that “T-theorizing fails to be *genuinely interpretative* for all languages where the full significance of the elements, gauged by their contributions to what is stated by the sentences which contain them, exceeds their truth-relevant significance, gauged by what they contribute to the truth or falsity of these sentences” (my emphasis, 15). Davidson takes Foster’s challenge head on by using radical interpretation to investigate what it is for a theory to *be* “interpretative.”

to understand what the natives *mean* by their words. Davidson is not resuscitating the idea that there is a *univocal* meaning for each sentence that the interpreter seeks to discover, which he takes Quine to have correctly eradicated.<sup>5</sup> Rather, his interpreter is trying to justifiably attribute the *intention to mean something* to the natives—a consideration that is peripheral and, strictly speaking, unnecessary, to Quine’s translator.

Davidson puts this point by saying that, strictly speaking, translation requires *three* languages, while interpretation requires only one (RI, 187). The aim of radical translation is the creation, in a *metalanguage*, of a translation manual which expresses systematic relations between sentences of the *object* (native) language and the *subject* (translator) language.<sup>6</sup> But the sole concern of radical interpretation is what natives mean by uttering *object* language sentences. So, the radical interpreter’s project is just an extreme version of our everyday interpretation of what other speakers mean when they utter *this* sentence in *this* context.<sup>7</sup> Davidson’s claim is only that the radical interpreter’s method of constructing a T-theory for L suffices for her to know what speakers of L mean; he is not committed to her being able to *express* that knowledge in any particular language (RF, 37).

---

<sup>5</sup> Davidson writes, “I don’t for a moment imagine...uniqueness would emerge [in interpretation]. But I do think that reasonable empirical constraints on the interpretation of T-sentences, plus the formal constraints, will leave enough invariant as between theories to allow us to say that a theory of truth captures the essential role of each sentence [in the language]” (RWR, 224).

<sup>6</sup> In principle, these three languages can be distinct. A person fluent only in the metalanguage can nevertheless use the radical translator’s manual to mechanically translate the statements of the translator and the natives, without having any idea about the meaningful content of their discussion.

<sup>7</sup> I think Davidson’s oft-quoted claim that “all understanding of the speech of another involves radical interpretation” is incautious (RI, 184). He does not think that we must tacitly or explicitly construct T-theories of each other to communicate (see RII, 127). This means critics who object that Davidson exemplifies philosophers’ tendency to erect *theories* where *description* is all that is wanted, such as Steven Mulhall (1987) and Richard Rorty (2000), attack a straw man. Rather, Davidson thinks that understanding others “involves” radical interpretation because we can justify our confidence in our everyday capacity *to* understand each other by reflecting on the conditions for successful interpretation that radical interpretation make explicit. It is a mistake (made by, e.g., Michael Root [1986]) to think that Davidson’s program turns on the controversial claim that radical interpretation “in the field” is possible. Instead, it turns on what he claims the radical species of interpretation can show us about the centrality of a shared concept of truth to everyday interpretation.

The radical interpreter begins her investigation by fallibly listing the contexts in which natives appear to *hold true* particular utterances, since an attitude of sincerity can (again, fallibly) be discerned prior to knowing what utterances are about (RI, 191-192). Of particular interest will be what Quine calls “observation sentences,” those sentences about which, on the basis of wide community agreement, it is possible to give a simple, plausible description of the circumstances in which natives hold them true. However, the meaning of the utterances are conflated with the natives’ beliefs in these held-true attitudes. For example, suppose the natives reliably hold true “gavagai” in the presence of rabbits. One day, a hare hops into view and the natives yell out “gavagai!” The radical interpreter is now faced with a choice. *Either* her initial hypothesis that “gavagai” is reckoned true by the natives if and only if they are in the presence of rabbits is faulty, and ought to be replaced in her list by a statement that “gavagai” is reckoned true if and only if they are in the presence of rabbits *or* hares, *or* her initial hypothesis about the meaning of “gavagai” was correct, but the natives erroneously believe that hares *are* rabbits. How is she to decide between these possibilities? Was she wrong about the *meaning* of the native *word* “gavagai,” or do the natives have a false *belief* about the *world*?

Davidson argues a charitable attitude is necessary to overcome this problem:

[We] solve the problem of the interdependence of belief and meaning by holding belief constant as far as possible while solving for meaning. This is accomplished by assigning truth conditions to alien sentences that make native speakers right when plausibly possible, according, of course, to our own view of what is right. (RI, 193)

So, Davidson would counsel the radical interpreter to suppose “gavagai” true in the presence of rabbits or hares. Beginning with her observation sentences, she should proceed to “optimize agreement” between her own beliefs and those of the natives (OVICS, 197). She seeks to replace each statement in her list with the form

The natives hold true “p” if and only if q

where “p” is a structural description of the native utterance and “q” is a description of features of the world, with a statement of the form:

“p” is true in the native language if and only if r

where “r” is either “q,” or a rephrasing of the observation expressed by “q” in terms of her own theory.<sup>8</sup> Equipped with her new list, the radical interpreter’s final goal is to isolate the component expressions of the native utterances (words) and analyze them (in terms of the features of the world to which they relate) in order to generate a theory that, for any native utterance, allows her to deduce a sentence from her list (which have the logical form of T-schema instances).<sup>9</sup> Her experimental attitude in constructing the list—since she will emend it to deal with any behavioural counterexamples that surface—gives her good empirical reason to believe that she has a T-theory for the native language that can be “projected to unobserved and counterfactual cases” (RF, 36). In short, according to Davidson, she can deduce from the fact that her T-theory meets the empirical constraints that she has a working theory of interpretation for the native language. She can now use her theory to attribute further propositional attitudes to the natives (about what they hope, desire, etc.) and engage them in meaningful conversation about the shared world.

Despite the similarities between the methods of radical interpretation and radical translation, it is worth stressing how rich the concept of truth becomes for Davidson in his quest to account for meaning.<sup>10</sup> For Davidson, it is part of the concept of truth, which is *sui generis*, that the evidence to which a theory of truth for a language must answer is the utterances that

---

<sup>8</sup> For instance, she might replace “the natives hold true ‘patooi’ if and only if they see the sky light up followed by a loud bang” with “‘patooi’ is true in the native language if and only if there is a thunderstorm nearby.”

<sup>9</sup> There is some room for trial and error here. If the interpreter cannot find a theory that perfectly entails the statements of her list, she may experiment with attributing some false beliefs to the natives.

<sup>10</sup> In chapter four, I shall argue that Quine’s and Davidson’s competing conceptions of truth is the heart of their dispute.



users of that language make about things in the shared world. We are only entitled to attribute a grasp of the concept of truth to a being who has a great many true beliefs about the world we share, and, inverting Tarski, the structure of the T-theory that allows us to so attribute truth also allows us to understand the language. Rendering the concept of truth intersubjective in this way is the heart of Davidson's humanism, and its central role in his arguments make it a point to which we shall have frequent occasion to return.

Davidson asserts that charity is not a choice, but a demand of radical interpretation:

if we cannot find a way to interpret the utterances and other behavior of a creature as revealing a set of beliefs largely consistent and true by our own standards, we have no reason to count that creature as rational, as having beliefs, or as saying anything. (RI 193; cf. CTTK, 234-235)

The objection to uncharitable interpretation is not that it evinces a mean-spirited imperialism, but that it is methodologically unsound. One reason to which Davidson appeals is that, far from making disagreement with the natives impossible, charity makes "meaningful disagreement possible" (MTM, 207), since "we can make sense of differences [in beliefs] only against a background of shared belief" (200). But critics have widely objected that this argument is inadequate. Anil Gupta, for instance, charges it with relying on a "plainly invalid transition" (2006, 195). Although it is plausible that interpreting natives requires *some* assumptions about their beliefs, Gupta denies that we are obliged to *charitably* assume a general agreement between our beliefs and theirs. We could assume, for example, that the natives only have correct beliefs during the daytime and interpret from there. But the more important reason that Davidson insists on charity is that finding meaning in the natives' utterances is the only way to find the natives' utterances *meant*. He thinks that we are only justified in attributing cognition to beings whom we interpret.<sup>11</sup> The eventual result of freely attributing false beliefs to the natives is not

---

<sup>11</sup> Nicholas Rescher objects that we are entitled to ascribe cognition and beliefs prior to interpretation if the beings whom we observe are "obviously" intelligent (1980, 328 ff.). He imagines encountering aliens who have constructed incredible architecture, but whose language we have yet to understand. But "obviousness" is surely too

concluding that they are foolish or irrational, but that they are *non-rational*, and that our impulse to interpret their squawks was misguided. So against Gupta, Davidson's point is not that charity is the *only* way to solve the problem of the interdependence of meaning and belief. He is arguing that charity is the only way to solve that problem while sincerely searching for a theory of meaning at all. Any other starting point risks judging some cognizers non-rational. Just as Quine before him, Davidson takes the radical linguist's basic assumption that the natives are speaking a language at all to be falsifiable on the basis of her continued observation. So Davidson insists that attributing false beliefs must be a last resort, only legitimate if we would otherwise be unable to develop a rational explanation for the natives' behavior.

This second, more important reason exposes the need to think of the requirement of charity in terms of justifying our ascription of the intersubjective concept of truth to the natives and finding them fellow cognizers, rather than in terms of optimizing agreement between us and finding them agreeably similar inquirers. In practice, charitably ascribing true beliefs to the natives will be indistinguishable from optimizing our agreement. Nevertheless, for Davidson, the force of the requirement stems from the fact that it is part of the concept of truth that the evidential basis for ascribing grasp of truth to a speaker are her utterances about the shared world, a fact which has the consequence that the beliefs of anyone to whom the concept of truth is correctly ascribed must be mostly true.<sup>12</sup>

---

treacherous a ground for our theories of mind. What initially appears to us as thoughtful architecture may, after all, turn out to be nothing but a complicated nest. Perhaps Rescher will counter that it is verificationist to conclude that only those beings whose cognition we can verify (having developed an interpretive theory for their utterances) are cognizers, but I shall delay discussing the mistake in charging Davidson with verificationism until section two.

<sup>12</sup> Davidson writes: "it is impossible for an interpreter to understand a speaker and at the same time discover the speaker to be largely wrong about the world. For the interpreter interprets sentences held true (which is not to be distinguished from beliefs) according to the events and objects in the outside world that cause the sentence to be true" (CTTK, 235). I shall discuss how this conception of interpretation grounds Davidson's answer to external world skepticism, particularly as that answer parallels but contrasts with Quine, in chapter four.

Davidson's understanding of the concept of truth also determines the central role of logic in radical interpretation. "Convention T and T-sentences," he writes, "provide the sole link between intuitively obvious truths about truth and formal semantics," (DCT, 66) doing so by employing logical resources to "reveal structure" in natural languages (MTM, 205). Just as an early goal for Quine's radical translator is finding a way to map her logic onto native utterances, Davidson, in a nod to Quine's canonical logic, says that the radical interpreter may need to "read...the logical structure of first-order quantification theory (plus identity) into the language" (RI, 193). But whereas Quine rules the translation of logic indeterminate (WO, 60-61; R, 314; 319), and the attribution of our canonical logic to the natives a pragmatically justified way to devise a workable translation manual for them, Davidson writes that, "the semantic constraint in my method forces quantificational structure on the language to be interpreted, which probably does not leave room for indeterminacy of logical form" (RI, 193 n16; c.f. IR, 228).

What semantic constraint is Davidson thinking of here? Radical interpretation is a matter of constructing a T-theory for a language L in use that meets the empirical demands necessary for it to count as a theory of meaning. A language in use must have been learned by its (finite) users, and Davidson argues that, since natural languages have an infinite number of well-formed expressions, we need to account for "how an infinite aptitude can be encompassed by finite accomplishments" (TMLL, 8). Accordingly, Davidson insists that an interpretive T-theory must not only be recursively axiomatizable, but finitely axiomatizable. This constraint gives us an explanation of how finite speakers are able to learn it and disallows T-theories that fail to bring out the structure of how sentences depend upon their components for meaning (SNL, 56).<sup>13</sup> Any interpretive T-theory for L must use a logic strong enough to express recursive clauses for L-

---

<sup>13</sup> A simple example is the case where ML contains L, and a T-theory for L is constructed whose axioms are all sentences of the form " 'S' is true if and only if S."

expressions. But why should these facts about the *interpreter's* T-theory require reading any particular logic *into* L or attributing grasp of a particular logic to its users?

One might suppose the reason to be that because the recursive clauses of an *interpretive* T-theory not only express observed regularities about the natives (e.g., for all native utterances “x” and “y,” natives hold true “x blick y” iff they hold true “x” or they hold true “y”), but also establish the *meaning* of the native utterances (e.g., justify the conclusion that for all native utterances “x” and “y,” “x blick y” *means that* x or y), we are forced to interpret others as committed to our own logical truths. For example, suppose that, on the basis of her observations, a radical interpreter judges

- 1) For all native utterances “x,” natives hold true “x blick pok x” iff they hold true “x” or they do not hold true “x”

She charitably transforms this sentence:

- 2) For all native utterances “x,” “x blick pok x” is true-in-L iff x or not x

But since for any utterance x, “x or not x” is a logical truth for the radical interpreter, the right hand side of the biconditional is true. So from 2), she can infer

- 3) For all native utterances x, “x blick pok x” is true-in-L

In the context of her interpretive theory, 2) becomes

- 4) For all native utterances x, “x blick pok x” means that x or not x

3) and 4) together imply that the law of the excluded middle holds for the native language. This argument can be run *mutatis mutandis* for any of our logical truths.

But the conclusion of this argument is far too strong. It leaves no room for disputes about logic. Extending this example, it would be impossible for an interpreter using classical logic to discover that the natives she observes are constructivists whose language only licenses intuitionist inferences. At best, intuitionists would be parochially interpreted as cautious

reasoners who fail to grasp the classical logical truths of their language. Instead, if the natives' inferences turn out to be cautious, bold, or otherwise unlike our own, we ought to even-handedly credit them with employing a language with a weaker, stronger, or otherwise deviant logic.<sup>14</sup> This respects our interpretive goal of finding what *they* mean by *their* utterances, rather than uncharitably foisting false beliefs about the logic of their language upon them. In turn, we may have to interpret their logical constants as behaving differently to our own.<sup>15</sup> But now we seem in danger of losing any logical constraint on interpretation. If I am right that Davidson believes beings using deviant logics are interpretable, in what sense does interpretation determine logical form?

One answer is that Davidson thinks a language's quantificational structure is the only way we have to understand its ontology (RQ, 81). So while different speakers might choose, on the basis of identical evidence, to endorse stronger or weaker inferences about the things that they talk about, if we are to interpret the natives as talking about *things* at all we are forced to read quantificational structure into their language and to attribute quantificational reasoning to

---

<sup>14</sup> Neil Tennant argues that to construct an interpretive T-theory, the radical interpreter need only find "analogues" in the native language for the introduction and elimination rules of her logical constants (1999, 72). He also proves that constructing a T-theory only demands expressive resources that both classical and intuitionist, and even *relevance* intuitionist, logicians can agree upon (1987). The upshot of these claims is that classical logicians, intuitionist logicians, and intuitionist relevance logicians can all interpret one another.

<sup>15</sup> But not *too* differently. Davidson's reason, as we shall see in a moment, is that all language users are interpreters, and since interpreting the infinite expressions of a learnable natural language requires a finitely axiomatizable recursive theory (and so, as shown by the structure of an interpretive T-theory, an expressively adequate set of logical constants), all language users have at their disposal an expressively adequate set of logical constants. (This set need not be classical; see the previous footnote.) Additionally, Simon Evnine persuasively argues that the development of operators playing the *approximate* roles of our truth functional operators is inevitable in any beings who possess the capacity to create theories (2001). If Evnine is right, reading into the native language operators that approximate our truth functional operators will be necessary if we are to reason with them about the world. This recalls Frege's attitude to the first species of logical alien. If we can view the practice of other beings as sensitive to at least *some* of the principles of inference, evidenced by their *approximate* grasp of our logical constants and the logical truths which they underwrite, then we can judge them reasoners and—if we learn to interpret their utterances—engage in reasoning with them. If the native practice is not sensitive to any of the principles of inference, like that of the resolute second species of logical alien, the beings are not reasoners.

them. A second (and more fundamental<sup>16</sup>) answer is that the only way that we know to construct a truth theory for a significant part of our own natural languages—which are the only natural languages we know—is to employ quantificational analysis.<sup>17</sup> This gives us good reason to expect that any truth theory that we will be able to construct for a natural language with similar expressive resources to our own will be quantificational. So if we are to interpret the natives as using a language in which truths can be stated, and so, as possessing the concept of truth, we must discover quantificational structure in their chains of reasoning. These considerations act as constraints on logical form: all interpretable languages are quantificational. Davidson’s further argument that all *languages* are quantificational is trickier; I will argue in section two that it does not presuppose verificationism.<sup>18</sup> For now, let us turn to the second claim which, combined with this view of the sense in which quantificational logic provides the necessary structure for interpretation, reveals Davidson’s conception of logic’s relationship to cognition.

Since it is uncontroversial that interpreting a person entitles us to ascribe beliefs about the world to her, and beliefs are necessary for cognition, one link between interpretation and

---

<sup>16</sup> Although detecting an ontology in the native language is important to the interpreter’s project of finding the natives’ utterances meaningful, Davidson writes of his essay “Reality without Reference” that “how a theory of truth maps non-sentential expressions on to objects is a matter of indifference as long as the conditions of truth are not affected” (1984, xix). Beyond demonstrating that Davidson follows Quine in the inscrutability of reference, this quote shows that the fundamental step in Davidson’s interpretation is developing a truth theory for the native language, which justifies attributing the concept of truth to the natives. Whereas Quine treats “truth” and “reference” even-handedly, taking both to be secured by the acceptance of homophonic instances of the paradigms for truth and reference by speakers of the home language and extended to other speakers via translation, Davidson takes truth to be more fundamental than reference for cognition and for the interpretation of others.

<sup>17</sup> Truth-functional analysis, for example, fails to capture our patterns of reasoning about sentences that are not truth-functionally compound, such as generalizations. But Davidson shows how various parts of natural language admit of quantificational analysis against first impressions, such as our reasoning about actions (LFAS, 105-110). Disciples of Davidson have continued this project, and the quantificational analysis of natural language is a live program in linguistics (see e.g. James Higginbotham, 2009).

<sup>18</sup> Lepore and Ludwig, for instance, judge correctly that Davidson thinks “[quantificational structure is] a constraint on interpretation forced on us...by our concept of language and truth,” but lament that Davidson does not make clear “what reason could be mustered for holding a priori that any language we wish to interpret will be *correctly interpretable* with the recursive resources of standard quantification theory” (2005, 215, my emphasis). In my view, Davidson’s humanism gives him a reason since, like Frege and Quine before him, he rejects the availability of the external standard of correctness that Lepore and Ludwig seek. In Davidson’s view, anything we can call a language is correctly interpretable by us, and we have good reason to think using a quantificational T-theory will allow us to interpret any natural language. I return to this point in section three.

cognition is clear. But as I mentioned above, Davidson additionally claims that we are *only* entitled to ascribe beliefs to a person if we can interpret her. So, he believes that we can extract from the conditions upon an interpretive theory the limits of other minds. In fact, the claim in which I am most interested, and which distinguishes Davidson's position from Frege's and Quine's, is stronger still. Davidson argues that we are only entitled to ascribe beliefs to those *who can interpret others*, indeed, to those who *have interpreted* others (TT, 156). This holds for the first person case as well as third person cases; insofar as we ascribe beliefs to ourselves, we must ourselves have interpreted other minds.<sup>19</sup> Davidson thus concludes that the capacity for cognition is resolutely interpersonal.

Here is a sketch of Davidson's argument for this claim:

- 1) Every cognizing being has beliefs.
  - 2) One possesses the concept of belief if and only if one possesses the concept of truth.
  - 3) Every being with beliefs has the concept of belief (and so, given 2), the concept of truth).
  - 4) Interpreting another person is the only way to acquire the concepts of belief and truth.
- So* Every cognizing being has interpreted another person.

Premise 1) is uncontroversial. To be capable of knowing, one must be capable of believing. The second premise is somewhat controversial. Davidson's argument for it is that "someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error—true belief and false belief" (TT, 170). Lepore and Ludwig object that empirical study has shown children as old as three "pass through a developmental stage in which they are unable to recognize that they have had false beliefs" (2005, 395).<sup>20</sup> But

---

<sup>19</sup> Pioneering sociologist George Herbert Mead's argument that a "self" can only arise through social interaction, since it is only through communication that an individual treats itself as an object ("one is talking to one's self as one would talk to another person" [1934, 140]) importantly prefigures Davidson's argument from triangulation. I plan to explore this connection further in later work.

<sup>20</sup> The studies they mention involve children, who, when faced with a candy box, naturally expressed their beliefs that the box contained candy (J. Perner, S.R. Leekam and H. Wimmer, 1987; A. Gopnik and J.W. Astington, 1988). Upon opening the box, the children found that it contained pencils. The children subsequently claimed not only that

Davidson can insist that this research merely shows some children as old as three, just as prelinguistic infants and animals, lack beliefs (RA, 95). If a being's utterances are to count as expressions of *beliefs*—as expressions of *our concept* of “belief”—they must sufficiently resemble the role played by those of our own utterances we take to express beliefs. If Davidson is right that our concept of belief “[stands] ready to take up the slack between objective truth and the held true” and that “we come to understand [belief] just in this connection” (TT, 170), then attributing “beliefs” to children who cannot grasp error is a mistake on a continuum with attributing beliefs about bananas to the monkeys who gesture at them or beliefs about daddy to the baby shrieking “Dada!”<sup>21</sup>

To argue for 3), Davidson first notes that we use the word “belief” when we wish to ascribe the endorsement of a particular propositional content to some mind. It is the paradigm of a propositional attitude. But “in order to have any propositional attitude at all, it is necessary...to have a belief about some belief” (RA, 104), because this is the only way to explain the surprise experienced by a fallible believer who discovers that the world is not correctly described by the propositional content which she previously endorsed. A surprised person is aware that although she believed *p*, she now believes *not-p*, and “such awareness...is a belief about a belief: if I am surprised, then among other things I come to believe that my original belief was false” (104).

---

they had *always* believed that the box contained pencils, but that other children who came into the room would *also* believe that the box contained pencils.

<sup>21</sup> One might object that this line is too hard. Denying that mute brutes believe is one thing, but denying our three-year old children believe is quite another. Can't we understand our children's utterances as the expressions of a primitive sort of belief? I think that for Davidson, this is already to concede the point he is trying to make. Certainly we *could* decide to call the utterances of three-year old children expressions of “proto-beliefs” because of their similarity to *some* features of actual fully-fledged beliefs. According to Davidson, doing so would be to employ a metaphor. In describing the behavior of children in terms of beliefs, what I say is literally false, but doing so allows me to articulate a striking feature of my world. As he puts it, “metaphors often make us notice aspects of things we did not notice before; no doubt they bring surprising analogies and similarities to our attention; they do provide a kind of lens or lattice...through which we view the relevant phenomena” (WMM, 1984, 261). Nevertheless, the helpfulness of metaphorically attributing “beliefs” to children is on a continuum with those used in naïve science: “someone might easily have no better or alternative way of explaining the movements of a heat seeking missile than to suppose the missile wanted to destroy an airplane and believed it could by moving in the way it was observed to move (RA, 101).



Since one can only have a belief about a belief if one has the concept of belief, Davidson concludes that all believers have the concept of belief.<sup>22</sup>

Hans Glock has recently criticized Davidson's description of surprise, arguing that he "has no argument to rule out the possibility of *a* simply switching from a belief that *p* to a belief that *q*, *without* that switch involving *a* believing that its original belief was false" (2007, 287, original emphasis). We might add that sometimes we are surprised because our new belief was unlikely on the basis of our current beliefs, not because we previously believed it false. But Davidson's point is not that *everything* we call surprise involves a belief about a belief, but rather that if a being lived its entire life without experiencing this sort of surprise, there would be no reason to call its attitudes "beliefs" at all. We would have no way to understand *a* as recognizing the difference between the world as it currently *appears* and the world as it in fact *is*, and it is only in the recognition of this difference that our concept of "belief" finds purchase.<sup>23</sup>

Premise 4) is the most controversial. Interpreting another person is surely *one* way that one might acquire the concepts of truth and belief. In struggling to understand what our interlocutors mean by their utterances, we may judge their view of the world simply wrong, and

---

<sup>22</sup> Some critics, such as Tomáš Marvan, mistakenly think that in the context of his denial of uninterpretable languages, Davidson's argument from surprise relies on the anthropocentric assumption that surprise can be reliably detected in all those whom we interpret (2003, 100ff). Marvan worries that we might be unable to interpret an alien language because we miss the aliens' expressions of surprise. But whether or not we could recognize alien surprise *as* surprise is irrelevant. Davidson's argument is just that if the aliens have beliefs, then they have the capacity for experiencing surprise. We may find it difficult to interpret aliens who express their beliefs differently. There is nothing unintelligible about a language that humans find hard to interpret. Davidson's claim that an uninterpretable language—a language that *cannot* be interpreted—is unintelligible is quite different, as we shall see in the next section.

<sup>23</sup> Similarly, Lepore and Ludwig's objection that an omniscient being has beliefs but cannot be surprised is unconvincing (2005, 395). They are right that an infallible omniscient being who knows the truth value of all propositions cannot be surprised in the sense of discovering that some content she had endorsed is false. But there would be no reason to describe this being as having "beliefs" in our sense of the word, precisely *because* there would be no possibility that what she held true was not in fact true.

But if an omniscient knower lacks the capacity for beliefs, ought we doubt the first premise that I called "uncontroversial"? I rather think that, on the contrary, Lepore and Ludwig's case calls into question the capacity for humans to conceptualize an omniscient mind. (Similarly, I do not think that Davidson's reliance on his learnability requirement to exclude infinitely recursively axiomatizable truth theories for interpretable languages begs any questions against infinite aliens, since we have no clear grasp of what such beings would be like.)

by reflecting on the norm underwriting this “wrong”—and simultaneously interpreting *ourselves*—we might realize that our own view of the world could similarly be mistaken. We grasp, that is, the potential for a gap between how we *believe* the world to be, and how the way the world *truly is*. But why should we think that interpreting another person is the *only* way in which the concepts of truth and belief can be acquired?

Many have found Davidson’s reason wanting: his report that he has “[no] idea” how “one could come to have the belief-truth contrast” other than through communication with another person (RA, 104). By relying on the limits of his imagination, Davidson leaves himself open to a critic who can articulate an alternative way of acquiring the twinned concepts of “belief” and “truth.” Lepore and Ludwig have recently taken up this challenge (2005, 402). They imagine an isolated being who has never communicated with another person looking back at a table at which she had previously glanced and thought she had seen a book, and now finding the table empty. Lepore and Ludwig argue that simplicity might lead the isolated being to conclude her past perceptual judgment was mistaken, rather than supposing that a mysterious disappearing book is in the vicinity. She would thus acquire the belief-truth contrast without help from another person.

Davidson introduces the term “triangulation” to describe the way that two communicators acquire the truth-belief contrast by acknowledging disagreements about the nature of the world that they mutually perceive. Lepore and Ludwig effectively claim that a single creature could rely on her memory to triangulate with her past self. But the triangulation Davidson thinks necessary for acquisition of our concept of “belief” involves more than merely finding three things to occupy its vertices. We use the word “belief” to ascribe endorsed propositional contents *about* the world to minds. If the cognitive processes of the beings in the triangle are to

count as beliefs, the triangle must somehow ground this relation of “aboutness.” Davidson argues his interpersonal triangle does so because the occupants of the “I” and “you” vertices fix the parts of the world that their beliefs are about by mutually perceiving them. He makes the same point in another way by insisting that the reference of the concepts possessed by a non-linguistic animal cannot be intelligibly located at any particular place in the casual chain from the world to its senses (SP, 119). So, against Lepore and Ludwig, if a person’s cognitive activity is to count as her believing things about objects in the external world, interpersonal triangulation (communication at a particular time with another person) is necessary. This insight about the way triangulation fixes the content of our beliefs buttresses Davidson’s assertion that communication is the only way that our belief-truth contrast can be grasped.<sup>24</sup>

Although I do not pretend to have provided a comprehensive defense of Davidson’s argument by showing how he can respond to some recent objections, I hope to have made his claim that interpretation is necessary for cognition somewhat plausible. Before turning in section two to the charge that Davidson’s views presuppose verificationism, let me conclude this section with a statement of his *humanist* conception of logic’s relationship to cognition (I shall defend the label “humanism” later):

**HUM:           Attributing the concept of truth to another mind using  
                  quantificational logic to structure the other mind’s language  
                  is a necessary condition for being minded.**

---

<sup>24</sup> Davidson is taking up a thread from Wittgenstein in insisting that a person can only be said to have a concept about an object if there is an answer to question of *which* object the concept is about. Here, as elsewhere, Davidson is not insisting that we as observers of Lepore and Ludwig’s self-triangulator—nor even the self-triangulator herself—must be capable of *verifying* which objects its concepts are about. Rather, he is claiming that the application of our concept of “belief” demands that there *be* an answer to this question, which requires *two* people observing the world, each of whom can act as a standard for the other. The self-triangulator only forms a *square* between her past self, her current self, the book seen by her past self, and the empty table seen by her current self. There is no principled way to isolate objects within the continuous stream of sensory information that affects her current self and her past self *about* which either could be interpreted as having beliefs (SP, 121).

This definition brings together Davidson's claim that logic is required for interpretation (in the sense that interpreting a learnable natural language requires using a finitely axiomatizable truth theory which reads quantificational structure into the interpreted language) and his claim that interpretation is required for cognition (in the sense that all believers have interpreted others). It is distinct from both Quine's naturalist and Frege's constitutivist conception of logic's relationship to cognition. Unlike Quine, who conceives of translating others as adopting our canonical logic as a *pragmatic* step justified by our desire to engage in inquiry with them, Davidson rules the use of logic to interpret others as fellow inquirers with a grasp of the concept of truth *necessary* for being an inquirer oneself. In this way, Davidson recovers the force of Frege's view that accepting the authority of logic is a prerequisite of cognition. But while Frege's constitutive view treats logical laws as norms of our inferential practice that any cognizer must acknowledge, Davidson articulates the conditions that must obtain for a cognizer to count as having acknowledged those norms, arguing that they can only be grasped interpersonally. We can think of it this way: if the solitary alien whom we have met on a distant planet can be understood as conforming to our principles of inference, Frege would rule her a fellow cognizer; Davidson would insist that she only *became* a cognizer through learning how to communicate with us, for only through triangulation with us did her cognitive activity count as her forming beliefs about the world. Frege's constitutivism pales in comparison with Davidson's nuanced account of the mutual emergence of the concepts of truth and belief in members of a community of cognizers who are capable of intersubjective inquiry into the objective world. Eschewing a theory which purports to explain epistemic concepts in terms of others, Davidson instead elaborates the conceptual interdependencies between the ingredients of epistemology.

In this section, I have argued that Davidson offers a third conception of the relationship of logic to cognition. More than once, matters turn on his claims about the correct application of truth and belief. Is Davidson's account of these concepts, and others central to epistemology, correct? Moreover, his argument that all languages we can interpret have quantificational structure only has bite in the context of his further argument that there are no languages that we *cannot* interpret, so that it follows from all cognizers being interpreters that all cognizers use the quantificational logic we recognize. A litany of philosophers think this argument can only go through on the assumption of verificationism.<sup>25</sup> In the next section, I shall engage Simon Blackburn's version of this objection, revealing why it mistakes Davidson's position. This work will allow me to articulate the commitments of Davidson's humanist conception of philosophy in section three, and show how those commitments underwrite his conceptual analyses of the core terms of epistemology.

## **2. The Charge of Verificationism and Davidson's Transcendental Argument**

Tarski is right, I think, in proposing that we think of natural languages as essentially intertranslatable. The proposal idealizes the flexibility and expandability of natural languages, but can be justified by a transcendental argument. (DCT, 72)

In this section, I shall argue that Blackburn's criticisms misunderstand Davidson's rejection of uninterpretable languages. By denying, with Quine, that our theories of reality are independent of our theories of how we know about reality, Davidson leaves no room for the charge that he is a verificationist who presupposes the knowability of what there is. The transcendental argument about interpretation that Davidson actually offers (and from which he

---

<sup>25</sup> See Blackburn (1984, 61), Dorit Bar-On (1994, 149), P.M.S. Hacker (1996, 306-307), Tennant (1999, 80), Gopal Sreenivasan (2001, 8-9), Marvan (2003, 95), Glock (2003, 198), Lepore and Ludwig (2005, 312-313), and Timothy Williamson (2008, 260). Even Nalini Bhushan, who correctly sees Davidson's transcendental aspirations, mistakenly accuses him of begging the critical question (1996, 253).

draws the denial of conceptual schemes as a consequence) depends upon his view that sharing the concept of truth is the cornerstone of the attribution of cognition.

Davidson's argument that all languages are interpretable is meant to show the sterility of conceptual scheme rhetoric.<sup>26</sup> Once liberated from such misleading vocabulary, epistemologists can identify the dualism between scheme and content as the "third dogma of empiricism" and acknowledge that we possess "unmediated touch with the familiar objects whose antics make our sentences and opinions true or false" (OVICS, 198). Although Davidson thinks that recognizing the third dogma sounds the death knell for empiricism, he does not herald rationalism the victor (189).<sup>27</sup> He seeks instead to overthrow the first-person Cartesian approach to epistemology that is shared by traditional empiricists and rationalists alike.

Davidson has no objection to the use of "conceptual scheme" as shorthand for the *unusual* way that we might find our interlocutor describing our world.<sup>28</sup> Rather, he objects to the rhetoric employed by thinkers like Thomas Kuhn, who claims that the difference between, for instance, a Newtonian and an Einsteinian, amounts to a difference of *world lived in* (1962, 134). Davidson takes Kuhn's idea to be that some of an Einsteinian's assertions simply cannot be understood by a Newtonian. The schemes are, as Kuhn says, "incommensurable," since in them the use of basic terms like "force" and "mass" differ (1970, 266-267). Hence, each scientist is not merely unusual but unintelligible from the other's perspective.

Davidson objects that the cogency of this "incommensurability" depends upon the cogency of a conceptual scheme being so different from our own that we cannot interpret it. The

---

<sup>26</sup> In fact, Davidson thinks conceptual scheme rhetoric is not merely sterile, but pernicious. He thinks it leads to epistemological skepticism (MTE, 55-56). I shall return to this point in chapter four.

<sup>27</sup> Despite endorsing *a priori* knowledge, rationalists have been just as guilty as empiricists of supposing there to be an intelligible distinction between scheme and content. I owe both Shawn Standefer and Thomas Ricketts thanks for helping me see this point.

<sup>28</sup> Desert-dwellers might have a dozen finely grained predicates for our own "sandstorm," for example. There is no obstacle to interpreting these various predicates by adding modifiers to the predicates of our own language (such as "mild," "moderate," and "severe").

language spoken by a person with such a scheme would have to be similarly uninterpretable.<sup>29</sup>

Davidson aims to show that this is not intelligible. *Ex hypothesi*, the vocabulary of an uninterpretable language cannot be systematically defined in terms of our own language. So some criterion is required by which we can identify a putative uninterpretable language L in use *as* a language *independently* of its relationship to our own language. Davidson proceeds to argue that the two families of metaphors philosophers employ in lieu of such a criterion, that “conceptual schemes (languages) either *organize* something, or they *fit* it” (OVICS, 191), are inadequate to the task.

To make sense of the uninterpretable language “organizing” something, there must be some plurality being organized. The likely contenders are *things* in the world, or *experiences* of L-speakers. But the case of L organizing things in the world reduces to that of an unusual, but interpretable, language. Although L-predicates may organize the world differently than English predicates, Davidson shows that a skilled interpreter fluent in either L or English can develop an interpretive T-theory for speakers of the other language and use it to judge the meaningfulness of their predicates. Suppose instead that L organizes an L-speaker’s experiences. Davidson argues that we are entitled to use the word “organize” only if we have some way to individuate natives’ experiences into a plurality. But this individuation—for example, judging that the world stimulates L-speakers’ sensory organs in such and such a way—provides the resources necessary for radical interpretation of their utterances (OVICS, 192-193). Once again, this case reduces to an unusual, but nevertheless interpretable, language.

---

<sup>29</sup> At least, it would have to be uninterpretable when the speaker is talking about her uninterpretable scheme. My focus (like Davidson’s) will be on totally uninterpretable languages, but Davidson also argues against the possibility of partially uninterpretable languages. Since interpretation demands charity, he says the difference between the conceptual schemes of speakers of L1 and L2, where some part of L1 cannot be translated into L2, is only a difference in degree and not in kind: “we improve the clarity and bite of declarations of difference, whether of scheme or opinion, by enlarging the basis of shared (translatable) language or of shared opinion. Indeed, no clear line between the cases can be made out... [W]hen others think differently from us, no general principle, or appeal to evidence, can force us to decide that the difference lies in our beliefs rather than in our concepts” (OVICS, 197).

Turning to the second metaphor, Davidson notes that empiricists claim a person's sensory stimulations constitute her sole evidence for holding particular sentences of her theory true, and that a theory "fits" the world if ongoing sense experiences square with its predictions. So, under empiricism, a theory is true just in case it successfully "fits" the total possible sensory evidence. But now, Davidson objects that, "the notion of fitting the totality of experience...adds nothing intelligible to the simple concept of being true" (OVICS, 193-194). The criterion that conceptual schemes must "fit" experience thus comes down to the claim that conceptual schemes must be (largely) true. Davidson argues that our best grasp on the desiderata for a theory of truth for an arbitrary L is Tarski's theory, and his T-schemata presuppose the notion of translation to explain "truth-in-L." Since to translate a language is to interpret it, we have made no progress. Davidson concludes that neither metaphor engenders a criterion for languagehood that can be applied to a practice independently of our capacity to interpret that practice in terms of our own language.

Davidson's final move is to argue that since we cannot make sense of a *radically* different conceptual scheme—only one that is unusual and perhaps difficult to interpret—we ought to jettison talk of conceptual schemes from epistemology. He thinks that, "it would be...wrong to announce the glorious news that all mankind...share a common scheme and ontology," for if we lack a criterion for saying conceptual schemes differ, we lack a criterion for saying conceptual schemes do not differ (OVICS, 198). We lack, that is, any way of making out a tenable dualism between individuals' uninterpreted experiences of the world and the "system" that they each employ to structure that experience. There is no epistemic intermediary between us and the world.



In *Spreading the Word*, Simon Blackburn voices an objection that has since become prevalent in Davidson scholarship:

[Davidson's] idea is that we can tell when a group is expressing ideas which are new to us only after we have become good at translating them in general. Davidson wants this point to sustain the conclusion that we can have no conception of what it might be for a group to have a language which expresses concepts and beliefs radically different from ours—too different to provide the backdrop of shared thoughts which make mutual translation generally possible. Such an argument relies on a principle of verification—an inference from the fact that we cannot *verify* that something is the case to the conclusion that we cannot conceive as a genuine possibility that it *might* be the case. (1984, 60-61, original emphasis)

Blackburn goes on to assert that there is now a “common consensus” that verificationism is dubitable.<sup>30</sup> Yet, even were Davidson to be granted his principle of verification, Blackburn claims, his argument fails. Although a *translator* whose own language lacks the requisite vocabulary may be unable to verify that her interlocutor has a conceptual scheme, one who takes the time to *learn* her interlocutor's language—in so doing, expanding her own vocabulary—can do so. (The conceptual scheme skeptic, Blackburn suggests, is like an ignorant listener who dismisses the Scottish bagpipes as noise.) Radically different conceptual schemes are perfectly intelligible once we acknowledge that we may need to acquire more vocabulary to verify their existence.

But Blackburn misreads Davidson.<sup>31</sup> Blackburn objects that an untranslatable language is intelligible because the object language may outstrip the expressive resources of the translator's subject language. Yet interpretation is not restricted to translation.<sup>32</sup> Davidson would agree with Blackburn about the intelligibility of there being one language that is

---

<sup>30</sup> This is the result of the perceived failure of logical positivism. A common objection is that the verification principle “all meaningful possibilities are verifiable” is itself unverifiable, and so meaningless by its own lights. I think that this objection vastly oversimplifies the guiding principles of different members of the Vienna Circle, and the extent to which logical positivism failed, but I cannot discuss this here. For recent contributions to this growing area of historical research, see Thomas Uebel (2004), Maria Carla Galavotti (2007), and Richard Creath (2011).

<sup>31</sup> I suspect that the misreading results from taking Davidson's paper in isolation, since in it he uses the terms “translation” and “interpretation” interchangeably (see e.g. OVICS, 196).

<sup>32</sup> As Quine would later put it, it is trivial that contemporary English is untranslatable in classical Arabic, since classical Arabic lacks the requisite vocabulary to express our concepts, but English is interpretable by speakers of classical Arabic (WWD, 75).

untranslatable in another. His objection is rather to the intelligibility of an *uninterpretable* language, one for which users of a natural language are *unable* to construct an interpretive truth theory. In his view, the sole reason we can have for calling some practice a *language* (or some cognitive process a *conceptual scheme*) is our capacity to interpret its *users* (RI, 184).

Nevertheless, the specter of verificationism still haunts Davidson's argument. Even granting that we can only verify some practice is a language if we can interpret its users, why should we be forced to conclude that there are no languages whose existence we cannot verify? If we join Blackburn and the "common consensus" in rejecting verifiability as a condition of possibility, is Davidson's argument undermined?

Davidson is well aware that one *could* appeal to verificationism to argue against conceptual schemes. But he says to do so—to directly infer from the fact that we cannot *verify* that a practice is a language that the practice is *not* a language—would be an uninteresting argument by fiat. If we take him at his word, the reason he unpacks the conceptual scheme metaphors at length is to *demonstrate* their emptiness, and so, to provide a positive argument against conceptual schemes based upon the meaningfulness of "truth" and "language."

Nevertheless, those critics sympathetic to Blackburn's objection may push it further and argue that, despite his protestations, verificationism underlies Davidson's dismissal of the second metaphor. Consider this passage:

[Convention T] isn't, of course, a definition of truth, and it doesn't hint that there is a single definition or theory that applies to languages generally. Nevertheless, Convention T suggests, though it cannot state, an important feature common to all the specialized concepts of truth. It succeeds in doing this by making essential use of the notion of translation into a language we know. Since Convention T embodies our best intuition as to how the concept of truth is used, there does not seem to be much hope for a test that a conceptual scheme is radically different from ours if that test depends on the assumption that we can divorce the notion of truth from that of translation. (OVICS, 194-195)

But, one might object, Convention T only requires that all sentences of L are translated in *some* ML in which the T-theory for L is constructed. Nowhere in Tarski can we find Davidson's

additional requirement that “we know” ML.<sup>33</sup> What prevents a T-theory for a language we cannot interpret being expressed in a metalanguage that we also cannot interpret? Since in this case we would not understand ML, we would not be able to *verify* that this T-theory for L was a T-theory. Yet to conclude from this that there are no uninterpretable languages seems to rely upon a principle of verification. Despite Davidson’s claim to have avoided arguing by fiat, has exposing this hidden premise shown he did just that?

I think that this objection, which I shall call the “hidden premise” objection, mistakes the character of Davidson’s argument. To see why, compare Martin Hollis’ earlier rejection of conceptual schemes:

If anthropology is to be possible...the natives must share our concepts of truth, coherence and rational interdependence of beliefs. Otherwise we are confronted as theorists with vicious circles. In other words Western rational thought is not just one species of rational thought nor rational thought just one species of thought. (1970, 218-219)

Hollis’ argument is explicitly conditional: *if* anthropology is to be possible—if, that is, we can develop theories and acquire knowledge of the native culture—*then* the natives must share our concepts. One may object that Hollis’ conclusion relies on the premise that anthropology *is* possible. Just because it may not be possible for Westerners to construct anthropological theories of a culture with a radically different, non-Western conceptual scheme, it does not follow that no non-Western cultures exist.

Returning to Davidson, the hidden premise objection supposes his argument similarly rests upon a conditional whose antecedent is a dubitable epistemological claim: *if we are to know* that some practice L is a language, *then* we must know an interpretive T-theory for L, and so know the language ML in which the T-theory for L is phrased. But, the objection continues, a practice may *be* a language even if we cannot *know* that it is a language. The only way to derive

---

<sup>33</sup> Bhushan, for example, objects: “Davidson is not justified in...simply taking it for granted that the only way to understand Convention T is in terms of translatability into a language we know” (1996, 253).

a negative existential from a conditional whose antecedent expresses a condition for our knowledge is to rely on an ultimately dubitable principle of verification.

But Davidson's argument against conceptual schemes rests upon no such conditional. According to the hidden premise objection, Davidson's argument turns on his claim that Convention T makes essential use of the notion of translation *into a language we know*. But this reading does not fit Davidson's text. Davidson writes:

Convention T suggests, though it cannot state, an important feature common to all the specialized concepts of truth. It succeeds in doing this by making essential use of the notion of translation into a language we know. (OVICS, 195)

If the reading grounding the hidden premise objection were correct, then Davidson ought instead to have written that *Convention T* or *It* makes essential use of the notion of translation into a language we know. His actual wording speaks rather to the *capacity* Convention T has to suggest something important about the specialized concepts of truth. So Davidson neither misunderstands the formal requirements of a theory satisfying Convention T, nor smuggles in an illicit connection between Convention T and languages we know. Instead, he claims that Convention T *is only able to suggest something important about the concept of truth* if we assume that the ML into which the sentences of L are translated is translatable into a language we know. Without this assumption, Tarski's logical machinery is idle. So Davidson's point is not only that we recognize the pertinence of Convention T to the concept of truth because of its application to our language; Convention T *is only* pertinent to the concept of truth so long as the connection to our language is preserved. Further, Davidson claims that Convention T embodies our "best intuition as to how the concept of truth is used" (205). So Davidson's objection to conceptual schemes is that it is contradictory to suppose a theory could be both largely true—i.e. expressed in an L whose ML T-theory is translatable into English—and uninterpretable.

Having exposed the irrelevance of the hidden premise objection, it is worth emphasizing the transcendental structure of Davidson's argument. He begins with the contingent empirical truth that we *do* interpret each other. That is to say, we find each other's utterances meaningful and meant. He proceeds to expound the conditions that must obtain for our interpretation of others to be possible—centrally, the justified attribution to our interlocutor of the concept of truth—and concludes that those conditions obtain. He analyzes core epistemic concepts like “truth” and “belief” by reflecting on interpretation because this is the process through which we find meaning in each other's utterances, and so, the activity that *makes* our words meaningful. In his view, “[Our] successful communication proves the existence of a shared, and largely true, view of the world” (MTM, 201), a conviction that drives Davidson's conclusion that we can make no sense of people having radically different conceptual schemes or speaking uninterpretable languages—thinkers or talkers who inquire about the world but whom we cannot understand.

On my reading, Davidson's argument rests upon two claims:

- 1) Convention T only plausibly expresses the conditions of a theory of truth for L on the assumption that the ML in which the T-theory for L is given is translatable into English.
- 2) Convention T is our best way of understanding the concept of truth.

There are two reasons for thinking that Convention T plausibly expresses the conditions for a truth theory of our own language. Firstly, disquotational English T-sentences for English (e.g. “snow is white” is true if and only if snow is white) are obvious. Secondly, their availability adequately explains our widespread practice of semantic ascent, one of the central uses of the word “true.”<sup>34</sup> These considerations explain Convention T's appeal *given* our intuitions about

---

<sup>34</sup> Furthermore, Davidson is optimistic that we can explain contexts in which we use the word “true” that do not straightforwardly relate to particular disquotational T-sentences (TM, 168-170). One sort of problem case, for example, is the use of “true” in the English sentence “Anything Amira says is true.” A Davidsonian might explain

how “truth” works in English. To evaluate 1), we must consider whether these considerations carry over to cases where we apply Convention T to other languages.

In the case of a *known* foreign language L, we can mention an L-sentence and ascribe truth-in-L to it just in those cases where we know it is accurately used. For example, knowing German, I can say: “Schnee ist weiss” is true-in-German if and only if snow is white. However, this T-theory for German would make little sense to a German who did not understand English. A T-theory that such a German could understand would need to be expressed in a ML that contained German (for example, “Schnee ist weiss” ist wahr genau dann, wenn Schnee weiss ist). A German-English bilingual understands that this T-sentence *is* part of a theory of *truth* because she understands German. A monolingual English speaker cannot. Does this sever the connection between Convention T and truth for monolinguals?

The answer is no, because, although the English monolingual does not understand German, she can interpret some of her fellow English speakers who do. She may acquire a justified belief that the T-sentence is part of a truth theory for German if a trusted bilingual tells her that it is. Additionally, the monolingual can grasp what it is for a sentence she does not understand to *be* a sentence of a truth theory for a foreign language because of hypothetical bilinguals who can provide a link between her grasp of truth and the meaningfulness of the foreign language.

However, the important case for evaluating 1) is when the ML in which the T-theory for L is expressed is uninterpretable to an English speaker. The proponent of the hidden premise objection tried to claim that the English speaker can grasp the possibility of what it is for an ML sentence to be part of a T-theory for L just by reference to her familiarity with T-theories, even if

---

this in terms of the speaker’s beliefs about the satisfaction of the conditions given by the T-statements of each of Amira’s future utterances.

ML is not interpretable. Such an English speaker would not, of course, be able to *tell* that a given ML sentence is part of a truth theory for L, but the critic insists that this epistemic claim has no bearing on the negative existential that there are no uninterpretable languages, unless a verification principle is somewhere operative. But Davidson's point, and I think he is right, is that in this case formal T-theorizing is simply irrelevant to the intersubjective concept of truth. Convention T plausibly expresses criteria for a truth predicate only so long as we preserve the theoretical link between our shared use of the word "true" and the sentences of the object language, so 1) is justified.

Let us turn to 2). Davidson's agreement with Quine that the analytic/synthetic distinction must be abandoned (and, hence, that there is no firm line separating the meaning of language from the truths of theory), acts as an important constraint upon potential ways he thinks that we can grasp truth. In his view, we can have no understanding of truth applying to theories that comes apart from our understanding of the truth predicate within our language. Although this constraint—and the need to reject the analytic/synthetic distinction—has been disputed, discussing objections to it here would take us too far afield.<sup>35</sup> Given this constraint, is Davidson right that Convention T is our best way of understanding the use of "truth"?

A number of critics think not. In discussing the second metaphor, Davidson writes that, "if a theory quantifies over physical objects, numbers, or sets, what it says about these entities is true provided the theory as a whole fits the sensory evidence" (OVICS, 193). But if this is right,

---

<sup>35</sup> Hacker, for example, complains that our ordinary conception of truth makes no mention of languages. He insists that "truth is not a property of sentences at all" (1996, 299), and that Davidson fails to recognize that "what is said (stated, asserted, claimed, etc) is that *p*, and it is this, not the sentence used in saying it, that is true or false" (300). For Hacker, it is not Convention T that captures our basic intuitions about "truth," but our agreement that, for example, it is true *that* grass is green if grass is, in fact, green. He thinks it an accidental fact that to convey this agreement we must make a statement in a particular language. But this view relies upon the coherence of propositions, of which truth is a property. I am unconvinced that Hacker satisfactorily answers the doubts Quine raises about the tenability of propositions, given that they lack clear individuation criteria, but space prevents me from further discussion here.

critics object, what licenses Davidson's subsequent focus on the sensory evidence available to *humans* in his account of truth?<sup>36</sup> Convention T makes truth predicates relative to particular languages, but they should also be relative to the sensory capabilities of particular languages' speakers. We should count a theory largely true *for a group of beings* if that theory fits the (total possible) sensory evidence *available to that group of beings*. Once we have made this concession, we can surely imagine a theory being largely true for a group of aliens, expressed in a language that is uninterpretable to humans, simply because our sensory experience of the world fails to coincide with the aliens' sensory experience of the world. Davidson's claim that Convention T is the best way of understanding truth, the critics assert, is (perhaps unwittingly) *anthropocentric*.

But, does this charge of anthropocentrism constitute an objection? Some defenders of Davidson, such as Marvan, have thought it a mere observation (2003, 105). On Marvan's modest reading, Davidson would admit that beings with radically different sensory capabilities can be said to have radically different conceptual schemes. Nevertheless, since humans share sensory capabilities, Convention T remains the best way of grasping the use of the concept of truth for beings like us.<sup>37</sup> The incommensurable *human* conceptual schemes invoked by the

---

<sup>36</sup> For versions of this objection, see Rescher (1980, 328ff), Hacker (1996, 300), Tennant (1999, 80), and Glock (2003, 198).

<sup>37</sup> But, one might object, not all humans *do* share sensory capabilities. Might the blind have different conceptual schemes to the sighted, or at least be incapable of interpreting color vocabulary? I am not sure what Davidson would say here. He acknowledges that the blind and the sighted can communicate (OQP, 191), but, to the extent that communication depends on interpretation, could the blind empirically confirm an interpretive T-theory for color vocabulary? Perhaps a blind person could use a device that transformed light frequencies into sound waves to identify the color of objects. But then, although it seems that she could construct a T-theory for a sighted person—"grass is green" is true-in-L if and only if my device makes a certain sound when pointed at grass—it is unclear that she could justifiably use this T-theory as a theory of meaning. After all, the sighted person does *not* mean that the device makes a certain sound.

On the other hand, Davidson might argue that by uttering "grass is green" in this context the sighted person *did* mean that the blind person's device makes a certain sound, *because* the utterance was directed to her: "The speaker wants to be understood, so he intends to speak in such a way that he will be interpreted in a certain way. In order to judge how he will be interpreted, he forms, or uses, a picture of the interpreter's readiness to interpret along certain



philosophers Davidson targets remain unintelligible. However, I think that Marvan's modest reading fails to do Davidson justice. The humanist backdrop of Davidson's transcendental argument permits a stronger response: namely, we lack any way to understand beings with radically different sensory capabilities possessing the intersubjective concept of truth that enables inquiry into the objective world, and so, as having cognition at all.

### **3. The Humanist Approach**

Aspects of our interactions with others and the world are partially constitutive of what we mean and think. There cannot be said to be a proof of this claim. Its plausibility depends on a conviction...*a priori* if you think, as I tend to, that this is part of what we mean when we talk of thinking and speaking. After all, the notions of speaking and thinking are ours. (CKVP, 294)

This quotation, taken from one of Davidson's final papers, speaks of his *a priori* conviction that interacting with others is "partially constitutive of what we mean and think." I think that apprehending the source of this conviction—that at which Davidson gropes by saying that "speaking" and "thinking" are notions that belong *to us*—is paramount if we are to understand him. This conviction forms the backdrop of his argument against uninterpretable languages, which is what gives force to his conception of the relationship between logic and cognition. Grasping Davidson's insight that our concepts belong to us reveals what I call his "humanist" conception of proper philosophical activity. The controversial analyses of "belief" and "truth" upon which his epistemological program depends are rooted in his commitment to our words only being meaningful in the context of our (human) interpersonal activity. Thus, humanist philosophers are committed to addressing problems by elaborating features of our *shared* inquiry.

---

lines" (NDE, 260). Perhaps Davidson would accept, as I think he should, the interpreter's "readiness *or capability*" as a friendly amendment to this claim, explaining the interpretability of the sighted by the blind.

Given Davidson's avowed anti-empiricism, and that he places no special weight on the *biological* category human, one may wonder why I am not content to call his position "rationalist."<sup>38</sup> But just as I see Quine's naturalism as *prior* to his commitment to empiricism, I see what I am calling Davidson's humanism as *prior* to a rationalist strain in his thought. Humanism drives his repeated appeals to "our" concepts and "our" language (RQ, 85; WQVT, 84). Here, we might helpfully contrast Davidson with Kant. Kant's practical and theoretical philosophy is intended to hold for any finite intellect equipped with the faculty of understanding and something like our faculty of sensibility (1999, A 230/B 283). His epistemic agents are abstract rational agents. But Davidson views mutual interpretation among humans based on our shared grasp of truth as the epistemic starting point. Our capacity to interpret each other, to attribute beliefs and desires to each other about the things in the world that we talk about together, is what makes us believers and desirers ourselves. Those whom we humans can interpret become—in that act—rational agents. Although it is unfortunate that the word "humanism" is pregnant with other meanings, it does serve to firmly locate Davidson's philosophy in reflections on *our* practice, which is more immediate to *us* than the practices of (our conception of) *rational beings*. The name also calls to mind the contrast between the human—or social—sciences and the natural sciences which are fundamental for Quine.<sup>39</sup>

From his humanist perspective, the major epistemic claim that Davidson seeks to defend is that an inquirer in possession of a radically constructed T-theory for another language has the capacity to interpret speakers of that language, in the sense of justifiably attributing the intersubjective concept of truth to those speakers, knowing something about their beliefs, desires,

---

<sup>38</sup> My thanks to Anil Gupta and Jamsheed Siyar for pressing me here.

<sup>39</sup> The potential alternative "socialism" is just as pregnant, and risks confusion of Davidson's externalism with Tyler Burge's *social* externalism. Unlike Burge, Davidson does not think that one's society is a relevant parameter for semantic theorizing (KOOM, 27).

and other propositional attitudes, and having the capacity to say what they are talking about. He thinks reflecting on this feature of radical interpretation will bring out the importance of the grasp of truth in everyday interpretation of other inquirers. In my view, just as Quine would relinquish empiricism but not naturalism if science demanded it, Davidson would relinquish his claim that we could necessarily construct a quantificational T-theory for any interpretable language but not his humanist insistence that the concepts we are investigating are generated through and depend upon intersubjective interpretation based on the shared concept of truth if he were presented with (for instance) undeniable speakers whose natural language proved quantificationally impenetrable.<sup>40</sup>

This principled humanism undermines the intelligibility of the anthropocentrism objection, which supposes that although we may grasp the concept of truth as a result of using our language with each other, our resulting grasp of the concept of truth can be applied independently of our ongoing practice. If this were right, then nothing would prevent us from hypothetically applying the concept of truth to other languages used by beings with radically different sensory capabilities, even though we would never be able to attribute grasp of truth to them. But Davidson's humanism denies the applicability of the concepts we use apart from their interpersonal, human, linguistic origin. The sole grasp we have of our concepts comes through our use of language, which is the sense in which our notions "belong to us," us human speakers. In this way, Davidson gives a humanist spin to Quine's rejection of cosmic exile. He believes

---

<sup>40</sup> Although we would be unable to interpret the speakers of this hypothetical language using a T-theory, Davidson would insist that they must be interpretable by us as possessing a grasp of the intersubjective concept of truth if they are rational beings at all. Davidson admits as a remote possibility that such a language could exist: "supposing it were clear (which it is not) that the deep structure of English (or another natural language) cannot be represented by a formal language with the usual quantificational structure, it still would not follow that there was no way of giving a theory of truth" (SNL, 58).

that there is no position outside of our ongoing inquiry which we could occupy and employ the concepts that only found expression within our language.

So, Davidson does not think it anthropocentric to think about the world humanly. There is no choice. To repeat, thinking and speaking are human concepts. To think or speak just *is* to think or speak humanly. As we saw in section one, we acquire the concept of truth—the truth/belief contrast—in virtue of speaking and interpreting other speakers of our language. We can extend a grasp of this concept to speakers of other languages only through interpretation. Davidson is unabashed about his focus on our practice, not because he only cares about conceptual schemes humans can have, but because “conceptual scheme” is, like all of our concepts, an ineluctably *human* concept.<sup>41</sup> Since Davidson views the primary application of “truth” to be contexts where we judge the world to accord with our beliefs, he concludes that “truths” *are* such as to be recognizable by humans, leaving no room for verificationist doubt that he conflates *truths* with *truths recognizable by those equipped with human sensory capacities*. (We might, perhaps, generate a new concept that allows the existence of truths we cannot recognize, but this is not the primary application of the intersubjective concept of truth with which we are concerned.<sup>42</sup>)

I think that we can see Davidson driving home a further consequence of his humanist rejection of uninterpretable languages in a late paper. Just as recognizing that there are no

---

<sup>41</sup> Tennant detects anthropocentrism in Davidson’s conception of “reality,” which he thinks the empiricist takes to refer to *all possible* sensory experience, not all possible *human* sensory experience (1999, 83-84). But, once again, Davidson can retort that we only grasp “reality” *through* interpreting each other as talking about the world. We can extend grasp of our concept of reality to others only by interpreting them, a process which has them acknowledging *our* reality *as* reality. As Bhushan colorfully puts Davidson’s attitude, “In the end, ‘language’ [and, I should add, ‘truth’, ‘reality,’ and *every* predicate] is an *English* predicate, not one for which God alone knows the correct extension” (1996, 261).

<sup>42</sup> This argumentative pattern should be familiar from chapter two. There, we saw Quine arguing against Stroud that the radical skeptic failed to grasp the meaningfulness of “reality” in our scientific discourse and that there was no non-philosophical use for a conception of “reality” capable of fuelling skeptical doubt. Here, Davidson argues that the critic pushing the anthropocentrism objection fails to grasp the meaningfulness of “truth” in our interpretive theories and that there is no use for a conception of “truth” capable of fuelling verificationist panic.

radically different conceptual schemes means recognizing that there are no conceptual schemes, recognizing that there are no uninterpretable languages means recognizing that

there is no such thing as a language, not if a language is anything like what many philosophers and linguists have supposed. There is therefore no such thing to be learned, mastered, or born with. We must give up the idea of a clearly defined shared structure which language-users acquire and then apply to cases. (NDE, 265)<sup>43</sup>

Davidson motivates this rejection by reflecting on our capacity to deal with malapropism. The title of Davidson's paper, "A Nice Derangement of Epitaphs," is a case in point. If sincerely uttered in conversation, a charitable listener will interpret this utterance to mean "A nice arrangement of epithets." Davidson notes that we regularly encounter speakers who use words in ways that we think are wrong, but rather than cutting short our conversation and staring at them in confusion, we are practiced in readjusting our theory of interpretation for them on the fly and continuing to communicate. This flexibility poses a problem for theories of language, like Carnap's, which suppose that competent users of a language L are able to communicate using it in virtue of sharing a more or less fixed interpretation of L's component expressions. Instead, Davidson argues, we need a theory of language use in which the "conventional meaning" of words for *every* speaker drops out.

His solution is once again grounded in the primacy of mutual interpretation for the humanist understanding of meaningfulness. We communicate with our interlocutors in terms of successively improving our *passing theories* of the meaning of their utterances, based on our *prior theory* of what we tacitly assume that particular words mean:

For the hearer, the prior theory expresses how he is prepared in advance to interpret an utterance of the speaker, while the passing theory is how he *does* interpret the utterance. For the speaker, the prior theory is what he *believes* the interpreter's prior theory to be, while his passing theory is the theory he *intends* the interpreter to use. (NDE, 260-261)

---

<sup>43</sup> It is worth noting that Davidson's late denial of languages does not undermine the learnability requirement he imposes upon interpreting languages, which is what forces interpretive truth theories to be finitely axiomatizable. Phrased in terms more conducive to the later Davidson, the learnability requirement states that we must be able to explain how any being we judge to have the capacity to communicate using an infinite variety of expressions developed that ability through finitely many steps.

“Language” is explanatorily inert in this account. It is grasp of the concept of truth through communication, not using language correctly according to some formal description, that is fundamental for mindedness. If we like, we can use the word “language” to refer to the passing theories of groups of speakers which cleave to one another, perhaps for historical, geographic, or cultural reasons. But it would be putting the cart before the horse to say that only those who speak the same language can communicate. Rather, only those who can communicate can be said to speak the same language.

Because of his humanism, and unlike Quine, Davidson is hostile to the possibility of finite natural languages. We can certainly design a T-theory for a language with finite expressions (as Tarski has shown us), but radical interpretation is meant to provide a T-theory that can be employed to detect *communicators* believing and acting for reasons, because the theory shows that they grasp the intersubjective concept of truth. Although each of us is a finite being, who use only a finite number of sentences in our lifetime, we all share the capacity as communicators to identify sentences drawn from an infinite range *as* sentences (for example, malapropisms or compound sentences that are recursively constructed using truth-functional operators on simple sentences). Putative speakers of a finite language would lack this capacity, meaning that we would be unable to recognize—to borrow Wittgenstein’s phrase—them as sharing our “form of life” (PI, 23; 241).

Let us take stock. I have argued that Frege, Quine, and Davidson agree that the logician seeks to codify the principles of inference that are implicit in *our* inferential practice. To do so, the logician must make ineliminable reference to our practice because we have no practice-independent grasp of valid inference. As Nelson Goodman puts it,

Principles of deductive inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences we

actually make and sanction...This looks flagrantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one...the process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either. (1983, 63-64)

All three philosophers agree with Goodman that the principles of deductive inference are *our* principles of deductive inference. Nevertheless, they develop this insight in importantly different ways. This results in their different conceptions of the relationship between logical and cognitive ability. Quine's naturalism, unlike Frege's constitutivism, does not claim that basic truths of our canonical logic express prerequisites necessarily acknowledged by all thinkers. Davidson reintroduces and transforms Frege's constitutivism socially, by making the concept of truth which our laws of logic unfold graspable only by *communicating* beings. Under humanism, the privileged relationship of logic to cognition is a *function* of logic's role in structuring truth theories.

This feature of humanism means that, unlike Frege and Quine, for Davidson the deviant patterns of reasoning exhibited by putative logical aliens is not the source of the problem. Whereas Frege and Quine both understand the patterns of inference that our logic codifies to be the base line for the attribution of cognition, Davidson takes the attribution of the intersubjective concept of truth to be fundamental.<sup>44</sup> In his view, beings with no grasp of truth are not fellow cognizers, and, since we have reason to think that beings who have a grasp of truth are interpretable using an interpretive T-theory that structures their language by imposing quantificational logic, logical aliens are ruled out.

---

<sup>44</sup> In contrast to Quine and Carnap, who he characterizes as interested in logic because they are "interested in improving on natural language," Davidson writes that he "[views] formal languages or canonical notations as devices for exploring the structure of natural language. We know how to give a theory of truth for the formal language; so if we also know how to transform the sentences of natural language systematically into sentences of the formal language, we would have a theory of truth for the natural language. From this point of view, standard formal languages are intermediate devices to assist us in treating natural languages as more complex formal languages" (MTM, 203). For Davidson, then, logic is in the service of truth.

Davidson's rejection of languages is strikingly consonant with Quine's rejection of meaning. Just as Quine argues meanings are unnecessary for explaining communication, Davidson jettisons fixed, formally specifiable languages in favor of dynamic interpretive theories between speakers. Both rejections are calls for philosophers to keep the subject matter they are attempting to explain firmly in view, and to avoid the temptation for extraneous theoretical assumptions that can distort rather than clarify. But despite this consonance, Davidson was convinced that his view was a substantive critique of, and improvement upon, Quine's philosophy of language. Quine disagreed. In the next chapter, I will sharpen the similarities and differences between their philosophical views by evaluating Davidson's objections, and suggest how we should resolve the debate between naturalism and humanism.



## **Bibliography**

Bar-On, Dorit. "Conceptual Relativism and Translation." *Language, Mind and Epistemology*. Eds. G. Preyer, F. Siebelt and A. Ulfig. London: Kluwer Academic Publishers, 1994. 145-170.

Bhushan, Nalini. "The Possibility of a Radically Different Language." *Philosophical Investigations*, 19 (1996): 237-263.

Blackburn, Simon. *Spreading the Word*. Oxford: Clarendon Press, 1984.

Burge, Tyler. "Individualism and the Mental." *Midwest Studies in Philosophy*. Eds. P.A. French, T.E. Uehling, and H.K. Wettstein. Minneapolis: University of Minnesota Press, 1979.

Creath, Richard, "Logical Empiricism." *The Stanford Encyclopedia of Philosophy* (Summer 2011 Edition). Ed. Edward N. Zalta. forthcoming URL = <http://plato.stanford.edu/archives/sum2011/entries/logical-empiricism/>.

de Caro, Mario, ed. *Interpretations and Causes: New Perspectives on Donald Davidson's Philosophy*. London: Kluwer Academic Press, 1999.

Davidson, Donald. "A Coherence Theory of Truth and Knowledge." (CTTK) *Kant oder Hegel?* Ed. Dieter Heinrich. Stuttgart: Klett-Cotta, 1983. 423-438. Rpt. in Davidson 2006. 225-237.

\_\_\_\_\_. "Comments on the Karlov Vary Papers." (CKVP) Kotatko et al 2001. 285-308.

\_\_\_\_\_. "In Defence of Convention T." (DCT) *Truth, Syntax and Modality*. Ed. Hugues Leblanc. Amsterdam: North Holland Publishing Co, 1973. Rpt. in Davidson 1984. 65-75.

\_\_\_\_\_. *Essays on Actions and Events*. Oxford: Oxford UP, 2001b

\_\_\_\_\_. *The Essential Davidson*. New York: Oxford UP, 2006.

\_\_\_\_\_. *Inquiries Into Truth and Interpretation*. Oxford: Clarendon Press, 1984.

\_\_\_\_\_. "Knowing One's Own Mind." (KOOM) *Proceedings and Addresses of the American Philosophical Association*, 1987, 441-58. Rpt. in Davidson 2001a. 15-38.

\_\_\_\_\_. "A Nice Derangement of Epitaphs." (NDE) *Philosophical Grounds of Rationality*. Eds. R. Grandy and R. Warner. Oxford: Oxford UP, 1986. Rpt. in Davidson 2006. 251-266.

\_\_\_\_\_. "Meaning, Truth and Evidence." (MTE) *Perspectives on Quine*. Eds. R. Barrett and R. Gibson. Oxford: Blackwell, 1990. Rpt. in Davidson 2005. 47-62.

\_\_\_\_\_. "Radical Interpretation." (RI) *Dialectica*, 27 (1973): 313-328. Rpt. in Davidson 2006. 184-195.

\_\_\_\_\_. "Radical Interpretation Interpreted." (RII) *Philosophical Perspectives*, 8 (1994): 121-128.

\_\_\_\_\_. "Rational Animals." (RA) *Dialectica*, 36 (1982): 318-327. Rpt. in Davidson 2001a. 95-106.

- \_\_\_\_\_. "Reality Without Reference." (RWR) *Dialectica*, 31 (1977): 247-253. Rpt. in Davidson 1984. 213-225.
- \_\_\_\_\_. "Reply to Foster." (RF) Evans and McDowell 1976. 33-41.
- \_\_\_\_\_. "Reply to W.V. Quine." (RQ) Hahn 1999. 80-87.
- \_\_\_\_\_. "Semantics for Natural Languages." (SNL) *Linguaggi nella Società e nella Tecnica*. Milan: Edizioni di Comunità, 1970. Rpt. in Davidson 1984, 55-64.
- \_\_\_\_\_. "The Inscrutability of Reference." (IR) *Southwestern Journal of Philosophy*, 10, (1979): 7-19. Rpt. Davidson 1984, 227-241.
- \_\_\_\_\_. "The Logical Form of Action Sentences" (LFAS) *The Logic of Decision and Action*. Ed. Nicholas Rescher. Pittsburgh: University of Pittsburgh Press, 1967. Rpt. Davidson 2001b. 105-148.
- \_\_\_\_\_. "On the Very Idea of a Conceptual Scheme." (OVICS) *Proceedings and Addresses of the American Philosophical Association*, 47 (1974): 5-20. Rpt. in Davidson 1984. 183-198.
- \_\_\_\_\_. "On Quine's Philosophy." (OQP) *Theoria*, 60, 3 (1994): 184-192.
- \_\_\_\_\_. "The Method of Truth in Metaphysics." (MTM) *Midwest Studies in Philosophy*, 2: *Studies in the Philosophy of Language*. Eds. P. French, T.E. Uehling, Jr., H.K. Wettstein. Morris: University of Minnesota, 1977. Rpt. in Davidson 1984. 199-214.
- \_\_\_\_\_. "The Second Person." (SP) *Midwest Studies in Philosophy*, 17. Eds. P. French, T.E. Uehling, and H. Wettstein. Indianapolis: University of Notre Dame Press, 1992. Rpt. in Davidson 2001a. 107-122.
- \_\_\_\_\_. *Subjective, Intersubjective, Objective*. Oxford: Oxford UP, 2001a.
- \_\_\_\_\_. "Theories of Meaning and Learnable Languages." (TMLL) *Proceedings of the International Congress for Logic, Methodology, and Philosophy of Science*. Ed. Y. Bar-Hillel. Amsterdam: North Holland Publishing Company, 1965. Rpt. in Davidson 1984. 3-15.
- \_\_\_\_\_. "Thought and Talk." (TT) *Mind and Language*. Ed. Samuel Guttenplan. Oxford: Oxford UP, 1975. Rpt. in Davidson 1984, 155-170.
- \_\_\_\_\_. "Truth and Meaning." (TM) *Synthese*, 17 (1967): 304-323. Rpt. in Davidson 2006. 155-170.
- \_\_\_\_\_. *Truth, Language and History*. Oxford: Clarendon Press, 2005.
- \_\_\_\_\_. "What is Quine's View of Truth?" (WQVT) *Inquiry*, 37 (1994): 437-440. Rpt. in Davidson 2005. 81-86.
- \_\_\_\_\_. "What Metaphors Mean." (WMM) *Critical Inquiry*, 5 (1978): 31-47. Rpt. in Davidson 2006. 209-224.
- Evans, Gareth, and John McDowell, eds. *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press, 1976.
- Evnine, Simon. *Donald Davidson*. Stanford: Stanford UP, 1991.
- \_\_\_\_\_. "The Universality of Logic: On the Connection Between Rationality and Logical Ability." *Mind*, 110, 438 (2001): 335-367.
- Foster, J. A. "Meaning and Truth Theory." Evans and McDowell 1976, 1-32.

- Galavotti, Maria Carla. "Confirmation, Probability, and Logical Empiricism" in *The Cambridge Companion to Logical Empiricism*. Ed. Alan Richardson and Thomas Uebel. Cambridge: Cambridge UP, 2007, 117-135.
- Goodman, Nelson. *Fact, Fiction, and Forecast*. 1956. 4<sup>th</sup> edition. Harvard: Harvard UP, 1983.
- Gopnik, A. and J.W. Astington. "Children's Understanding of Representational Change and Its Relation to the Understanding of False Belief and the Appearance-Reality Distinction." *Child Development*, 59, 1988, 26-37.
- Glock, Hans. *Quine and Davidson on Language, Thought and Reality*. Cambridge: Cambridge UP, 2003.
- Gupta, Anil. *Empiricism and Experience*. Oxford: Oxford UP, 2006.
- Hacker, P.M.S. "On Davidson's Idea of a Conceptual Scheme." *The Philosophical Quarterly*, 46, 184, (1996): 289-307.
- Hahn, Lewis, ed. *The Philosophy of Donald Davidson*. Chicago: Open Court, 1999.
- Higginbotham, James. *Tense, Aspect, and Indexicality*. Oxford: Oxford UP, 2009.
- Hollis, Martin. "The Limits of Irrationality." *Rationality*. Ed. Bryan Wilson. New York: Harper and Row, 1970. 214-220.
- Kant, Immanuel. *Critique of Pure Reason*. [1781/1787] Trans and Eds. P. Guyer and A. Wood. Cambridge: Cambridge UP, 1999.
- Kotatko, P., P. Pagin and G. Segal eds, *Interpreting Davidson*. Stanford: CSLI, 2001.
- Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: Chicago UP, 1962.  
 ———. "Reflections on my Critics." *Criticism and the Growth of Knowledge*. Eds. I. Lakatos and A. Musgrave. Cambridge: Cambridge UP, 1970.
- LePore, Ernest and Kirk Ludwig. *Davidson: Meaning, Truth, Language and Reality*. Oxford: Oxford UP, 2005.
- LePore, Ernest, and Brian McLaughlin, eds. *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell, 1988.
- LePore, Ernest, ed. *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell, 1986.
- Ludwig, Kirk, ed. *Donald Davidson*. Cambridge: Cambridge UP, 2003.

- McDowell, John. "Functionalism and Anomalous Monism." LePore and McLaughlin 1988. 387-398.
- Marvan, Tomáš. "Interpretability, Perceptual Schemes and Triangulation." *Sats – Nordic Journal of Philosophy*, 4, 2, 2003, 93-107.
- Mead, George Herbert. *Mind, Self, and Society*. Chicago: University of Chicago Press, 1934.
- Mulhall, Stephen. "Davidson on Interpretation and Understanding." *The Philosophical Quarterly*, 37, 148, 1987. 319-322.
- Perner, J., S.R. Leekham and H. Wimmer. "Three year-olds' Difficulty with False Belief: The Case for Conceptual Deficit." *British Journal of Developmental Psychology*, 5, 1987. 125-137.
- Quine, W.V. *From Stimulus to Science*. Cambridge: Harvard UP, 1995.  
 ———. *Pursuit of Truth*. (PT) Rev ed. Cambridge: Harvard UP, 1992.  
 ———. "Replies." (R) *Words and Objections: Essays on the Work of W.V. Quine*. Eds. D. Davidson and J. Hintikka. Holland: D. Reidel, 1975, 292-352.  
 ———. *The Roots of Reference*. (RR) LaSalle, Illinois: Open Court, 1973.  
 ———. "Where Do We Disagree?" (WWD) Hahn, 1999. 73-79.  
 ———. *Word and Object*. (WO) Harvard: MIT Press, 1960.
- Ramberg, Björn. "What Davidson Said to the Skeptic or: Anti-Representationalism, Triangulation, and the Naturalization of the Subjective." Kotatko et al 2001, 213-236.
- Rescher, Nicholas. "Conceptual Schemes." *Midwest Studies in Philosophy V*. Eds. P. French, T. Uehling, and H. Wettstein. Minneapolis: University of Minnesota Press, 1980. 323-346.
- Root, Michael. "Davidson and Social Science." LePore 1986, 272-306.
- Rorty, Richard. "Response to Davidson." *Rorty and His Critics*. Ed. Robert Brandom. Oxford: Blackwell, 2000, 74-80.
- Sreenivasan, Gopal. "Understanding Alien Morals." *Philosophy and Phenomenological Research*, 62, 1, (2001): 1-32.
- Tarski, Alfred. *Logic, Semantics, Metamathematics*. Trans. J. H. Woodger. Oxford: Clarendon Press, 1956.
- Tennant, Neil. *Anti-Realism and Logic: Truth as Eternal*. Oxford: Oxford UP, 1987.  
 ———. "Radical Interpretation, Logic, and Conceptual Schemes." de Caro 1999, 71-94.
- Uebel, Thomas. "Carnap, the Left Vienna Circle, and the Neopositivist Critique of Metaphysics" in *Carnap Brought Home: The View from Jena*. Ed. S. Awodey and C. Klein. Chicago: Open Court, 2004, 247-278.

Williamson, Timothy. *The Philosophy of Philosophy*. Oxford: Blackwell, 2008.

Wittgenstein, Ludwig. *Philosophical Investigations*. (PI) 1951. Trans. G.E.M. Anscombe.  
Oxford: Basil Blackwell, 2001.

## **Disagreement in Philosophy: The Case of Naturalism Versus Humanism**

Readers of Davidson and me are bound to be struck by how deeply we agree, and hence puzzled the more by occasional points of *apparent* disagreement.

W.V. Quine (my emphasis)

Donald Davidson admits to being perplexed by W.V. Quine's paper "Where Do We Disagree?". Their views concerning the role played by sensory stimulations in translation "have converged" (RQ, 82), but he "[does] not see how to reconcile" Quine's empiricist epistemology with his own externalist semantics (84). Although Quine takes their remaining differences to be little more than a matter of emphasis, Davidson suspects that their disagreement is substantive.<sup>1</sup>

There is little critical consensus about how to resolve disagreement between Quine and Davidson.<sup>2</sup> They undoubtedly agree about a number of controversial claims in the philosophy of language and mind: both reject the analytic/synthetic distinction, both accept anomalous monism, and both think translation and reference are indeterminate. The points at which they part—such as how to define observation sentences, and the extent to which logical form is indeterminate—seem technical (perhaps even peripheral) in contrast. This has tempted some commentators to concur with Quine and to conclude their "disagreement" is *only* apparent.<sup>3</sup> If they agree upon all the essentials, we can isolate the shared features of their philosophical framework and criticize them together. Yet, I think that this approach deflects our attention from the most puzzling, and philosophically important, aspect of their discussion: what makes their differences seem incidental from Quine's perspective and fundamental from Davidson's?

---

<sup>1</sup> Of the three divergences with Davidson that Quine identifies, he casts their disputes about ontological relativity and observation sentences as essentially terminological (WWD, 73, 74), and suggests that his own views on truth are "closer to Davidson's than he seems to have thought" (77). Davidson denies the first two claims (RQ, 81-82), and is puzzled about how to evaluate the third (85), which I unpack below in some detail.

<sup>2</sup> Each have advocates who, at least, unite in finding them different. For a defense of Quine, see Dirk Koppelberg (1990) and Lars Bergström (2001); for a defense of Davidson, see Richard Rorty (1991) and John McDowell (1994).

<sup>3</sup> See, for example, Folke Tersman (2001). Others, such as George Romanos, find their differences of "very little philosophical interest" (1983, 183).

In this chapter I will diagnose why Quine and Davidson disagree about their disagreement. I will argue that this second-order disagreement stems from their differing naturalist and humanist conceptions of philosophy. Although Quine finds “naturalizing” the substance of Davidson’s philosophy conducive to his program, Davidson finds no philosophical value in “humanizing” Quine’s work, resulting in opposed evaluations of their intellectual relationship. To suppress their differences is to obliterate the subtleties distinguishing their conceptions of philosophy and the corresponding limits they place on intelligibility.

In section one, I argue that Davidson’s claims that naturalized epistemologists cannot respond to skepticism about the external world and that naturalists mistakenly admit unconceptualized reasons fail to engage Quine’s view. In section two, I argue that Davidson’s real objection to Quine concerns the nature of truth and objectivity under naturalism. Although articulating their naturalist and humanist commitments serves to differentiate Quine and Davidson, it does not adjudicate their dispute. In section three (contributing to recent scholarly interest in the epistemology of disagreement), I use their dispute to examine options for resolving philosophical disagreement. I tentatively endorse a version of tolerance which, by forcing us to occupy that perspective which our current inquiry demands, avoids the fragmentation of philosophy into distinct totalizing programs.

### **1. Davidson’s Skeptical Attacks Against Quine**

Davidson’s criticisms of Quine range from incisive to misguided. In this section, I want to clear away those of his objections which misfire by exposing their various misunderstandings of, or lack of force against, Quine’s naturalism. This will allow me in section two to distill the

core of Davidson's objection to Quine's epistemology, which, I contend, concerns their different conceptions of truth and objectivity.

Davidson initially names Quine one of the philosophers in the grip of the third dogma of empiricism (OVICS, 191). But although Davidson quotes him using the phrase "conceptual scheme," Quine does not subscribe to the crude empiricist view that an individual's unconceptualized sensory experience is the evidence for her theories. He recognizes that we typically know nothing about our stimulations, and rarely if ever cite stimulations as the evidence justifying our beliefs. Rather, in developing his naturalist account of our epistemic and linguistic capacities, Quine follows current science in taking the access that beings like us have to the world to be sensory. As a physicalist, Davidson similarly believes that by impinging on our sense organs the world causes us to have beliefs about it. The dualism to which Davidson objects is thus not straightforwardly part of Quine's view.

Quine intends empiricism, as he puts it in his initial response to Davidson, to be a theory of warranted belief, not truth (OVITD, 39). He had explicitly rejected Charles Sanders Peirce's pragmatist definition of truth in terms of idealized warrant for belief (WO, 23-24), and if *that* was the source of Davidson's anxiety about a third dogma, Quine happily concurs with his conclusion: "what Davidson imputes to [empiricism is then] rightly imputed and rightly renounced" (OVITD, 39).<sup>4</sup> Nevertheless, naturalism demands philosophers align themselves with scientists. Given that our best scientific theories take us to be physical parts of the physical world, epistemologists ought to hold that we are warranted in believing those theories which best

---

<sup>4</sup> Quine thinks that we cannot define truth by equating it with the ideal scientific theory of the world, not just because our theories are underdetermined by the total possible evidence we could have for them (thus making it doubtful that there *is* a single ideal scientific theory), but also because of semantic holism. We cannot hold the sentences of our current theory to be true if and only if they "match" sentences within the ideal scientific theory, because isolated sentences are "meaningless intertheoretically" (OVITD, 24). Quine concludes that seeking an account of truth that goes beyond Tarski's analysis is idle. In section two, I will examine the deeper significance that Davidson awards the concept of truth.



explain and predict our sensory experience. If empiricism is so understood (as an account of the conditions to be met by any theory in which we are justified in believing) Quine sees no problem with describing the different—and increasingly sophisticated—scientific theories that we have developed as “schemes” employed to better predict and understand the world.

In “Meaning, Truth, and Evidence,” Davidson turns to skepticism to try and explain once again why scheme/content dualism is not an innocent feature of any epistemology, even one that (like Quine’s) conceives of empiricism as a theory of warranted belief. He acknowledges that stimulations do not play the role of “evidence” in Quine’s theory of evidence (MTE, 52). But, since Quine uses his behavioristic ersatz of meaning—“stimulus meaning”—to explain our construction of meaningful scientific theories (WO, 12), he appears to think stimulations play a *semantic* role.<sup>5</sup> Because of his emphasis upon an individual’s sensory organs, Davidson coins the term *proximal* to describe Quine’s conception of the way that stimulations are semantically relevant, in contrast to his own *distal* view which makes the world itself the sole semantically relevant stimulation.<sup>6</sup> Since naturalized epistemologists seek to elucidate the evidential support that each individual has for her theory, and make ineliminable reference to private, proximal stimulations in doing so, Davidson finds a problematic scheme/content strain in Quine’s thought. According to Davidson, this strain precludes naturalized epistemologists from accounting for our interpersonal knowledge of the world.

Davidson begins his argument by instructing us to imagine an interlocutor stricken with a “rearranged sensorium” (MTE, 55). This causes her to have rabbit-stimulations instead of

---

<sup>5</sup> For present purposes, we can define the stimulus meaning of an utterance *u* for a speaker *S* as the ordered pair of sets of stimulations that would elicit assent and dissent respectively from *S* if asked “*u*?” (Quine discusses some complications in WO, 32-34.)

<sup>6</sup> Davidson acknowledges, however, that elsewhere Quine seems to defend a distal view. In discussing ostensive learning, Quine argues that objects in the world, perceivable by both speakers, are the relevant locus for determining meaning: “The child and the parent must both see red when the child learns “red,” and one of them must see also that the other sees red at that time” (RR, 10).

warthog-stimulations whenever she is near warthogs. He points out that *if we know* of our interlocutor that she is being rabbit-stimulated in warthog situations, Quine's theory, by making private stimulations semantically relevant, tells us that we should translate her utterances in warthog situations to concern rabbits. In turn, we must conclude that our interlocutor is systematically mistaken about the world. But, having now appreciated that the "warrant" naturalism accords our beliefs is based solely upon how well our scientific theory predicts *our* stimulations, we must accept that *we* might be radically mistaken about the public world. Our own stimulations may be just as defective in presenting us the environment:

Although each speaker may be content that his view is the true one, since it squares with all his stimulations, once he notices how globally mistaken others are, and why, it is hard to think why he would not wonder whether *he* had it right. Then he might wonder what it could mean to get it right. (MTE, 55-56)

The last sentence suggests that Davidson is applying pressure to the account of normativity allowed by Quine's naturalism. His skeptical argument should not only undermine the naturalist's confidence that she knows about the world; he thinks it should make her question the very intelligibility on her view of knowing about—of being "right" about—the world. Davidson claims the only way to deny the skeptic this purchase upon epistemology is to reject the third dogma.<sup>7</sup> The only "stimulation" relevant to our knowledge of the public world is the world itself, with which humans thus possess *unmediated* contact (OVICS, 198).

However, as I argued in chapter two, Quine's naturalism has the capacity to rebut numerous skeptical attacks. There, I also showed how a Quinean could respond to Ebbs' doubts

---

<sup>7</sup> Nalini Bhushan misses this point (1996). She claims that, despite Davidson's argument against conceptual schemes, we can say that a *single* human conceptual scheme exists with which all humans organize their sensory experience. But in Davidson's view, single conceptual scheme epistemology falls to skepticism. Since it leaves the dualism intact, the skeptic can cogently doubt her putative knowledge on the basis that she might be a *deviant* human—or the next evolutionary step—whose internal wiring presents the world differently to her than to other human inquirers.

about the naturalist's account of normativity. Can Quine similarly disarm the two strands of Davidson's warthog argument?

Quine admits that the warthog argument demonstrates a problem with characterizing observation sentences as sentences whose stimulus meaning is approximately shared by members of a group. The radical translator begins constructing her manual by matching observation sentences in the native language to sentences in her own language. But the particular ways in which individual natives are stimulated should have no bearing upon the semantic question of how to translate their utterances. For translation, all that matters is detecting others reliably using the same expressions in the same contexts, which we can then translate with expressions *we* reliably use in those contexts. Davidson's argument depends upon "imagining someone who, when a warthog trots by, has just the patterns of stimulation I have when there's a rabbit in view" (MTE, 55). Quine's solution is to admit that people with unique neurophysiologies cannot "share" a stimulus-meaning. Another person can never have "just the pattern of stimulation" which I do. Removing "intersubjective likeness" of physical stimulation from his account, Quine now argues that intersubjectivity comes only with language use.<sup>8</sup> Choosing a manual which translates the warthog deviant as talking with us about rabbits is the best way to promote our ongoing communication.

Quine can now appeal to naturalism to defuse Davidson's skeptical objection in much the same way that he defuses Stroud's. He would ask Davidson to fill out the details of his challenge. If we have empirically discovered that some people have deviant warthog neurophysiologies—perhaps when they are wired to brain scanners, different parts of their brain

---

<sup>8</sup> Quine writes, "what floats in the open air is our common language, which each of us is free to internalize in his peculiar neural way" (PT, 44). He now defines a sentence as observational if it has a constant stimulus meaning for an individual over time, and observational for a community if, firstly, it is observational for each member, and secondly, if the members assent to or dissent from the sentence in the same observable circumstances.

light up when they see warthogs compared to the parts of the brain that light up in most of the population—and the skeptic is a *constructive* skeptic worried that she might be a deviant, the naturalist response is that she should get her brain scanned. If the point is instead that undiscovered deviant neurophysiologies might exist, the naturalist will counsel this *pessimistic* skeptic that, although our neurophysiological theories are fallible and doubtless incomplete, scientific realism is currently our best theory of reality. If the skeptic turns *radical*, and suggests that science may *never* reveal the “true” nature of “reality” to us, we ought to rebuke her for allowing her imagination to overrun her knowledge of semantics. She has failed to describe a meaningful possibility in the terms of our language.

It is worth taking a moment to observe the similarity between Quine’s views here and Davidson’s own response to skepticism about the external world, since the similarity is one reason that scholars have thought Quine and Davidson to be in fundamental agreement. Davidson thinks that observing the principle of charity is necessary when interpreting another speaker. To identify another person as having a mind, as being an entity capable of believing, we must attribute truth to those beliefs we interpret the mind having in terms of the world in which we both live. Davidson insists that a parallel act of interpretation occurs if we reflect upon our own beliefs. Determining what it means for me to so much as *have* a belief involves isolating my beliefs and ascribing truth to them in terms of the shared world in which I live with my linguistic community (CTTK, 236). So it is incoherent to worry that all of our beliefs might be false, because, in the first instance, our concept “belief” is correctly attributed to a mind *only* if most of the things we call its beliefs are true. Davidson, like Quine, charges the radical skeptic with having made a semantic error.

Were Davidson's imputation of the third dogma to Quine essentially a challenge to refute skepticism, then it is clear from Quine's perspective that it fails to breach his naturalism. But although Davidson thinks skepticism an inevitable consequence of any epistemology, such as Quine's, which refuses to jettison the third dogma, he insists that:

[the] central argument for rejecting empiricisms that base knowledge on something unconceptualized is that nothing that is unconceptualized can serve as a reason for an empirical belief, or for anything else, since the relation of a reason to what it supports is conceptual and so demands that the reason have a propositional content. (CKVP, 286)

Davidson thinks one can only justify one's knowledge by appealing to one's other beliefs, not by appealing to physical facts, such as how one's sense organs are being stimulated. Of course, one's *beliefs* about how one is being stimulated can be a reason for one's other beliefs—but in Davidson's view, this is just to accept that everything relevant to epistemology is conceptualized, and thus, that one ought to reject scheme/content dualism.

But, once again, Quine need not be compelled. The quarrel now seems to amount to terminological wrangling over the word "epistemology."<sup>9</sup> What is the central concern of epistemology? Quine's naturalized epistemologists seek "the rationale of reification" (TI, 3): a scientific explanation of how it is that we develop knowledge of the world, where science tells us that "we" are physical organisms equipped with five senses, and our experience is vibrant, conceptually complex, and of a world composed of diverse objects. Since our scientific theories hold that the physical world (of which we are a part) is governed by causal laws, our epistemological account ought to explain how the world *causes* us to have beliefs about it. Naturalism demands that we look to science to find our epistemic starting point. Current science returns with the answer: our sensory stimulations. The philosophical objection Davidson levels against this account, that unconceptualized facts *cannot* be "reasons" for beliefs, strikes Quine as

---

<sup>9</sup> Because of Quine's predilection for scientific methodology and his attention to neurology, Davidson claims that Quine is a biologist, while he is the "true" epistemologist (RQ, 83). Quine insists that he is the "true" epistemologist because he is concerned with our knowledge of the external world, labelling Davidson a semanticist (WWD, 74).

belying a retrograde explanatory stance.<sup>10</sup> Quine agrees that our personal experience of justifying our beliefs involves appealing to our other beliefs. Nevertheless, he thinks tailoring our theory of knowledge to our personal experience is to privilege introspection over scientific method, and so, to overlook the lessons we ought to learn from empirical investigation.

Davidson's theoretical inroads into our use of mentalistic predicates (which Quine agrees are indispensable for everyday discourse) and the dependence of thought upon language use can easily be integrated into Quine's naturalist program (PT, 72-73).<sup>11</sup> Davidson agrees with Quine that the physical world stimulates human organisms, which eventually results in individuals knowing about the world. Quine takes himself to be sketching an account of this entire causal story. He thinks Davidson is focused on (and produces a number of insights concerning) the later, social stages, in which inquirers learn how to gainfully respond to each other's speech. In this way, Quine thinks he can "naturalize" the parts of Davidson's work with which he agrees (anomalous monism, for example), and takes their viewpoints to be fundamentally consonant.

Quine unrepentantly continues to talk about various "conceptual schemes" of reality in subsequent articles (SN, 406; NLWOM, 471). But Davidson disagrees with Quine's evaluation of their relationship, remaining opposed to the naturalist account of epistemological justification. He is adamant that "an appeal to naturalism will not help" (MTE, 54) in satisfactorily defusing the third dogma he identifies. Why?

The reason is that neither skepticism nor the possibility of unconceptualized reasons gets to the heart of Davidson's misgivings about naturalism. Earlier, I noted that a second strand in his warthog argument concerned normativity. This gestures towards Davidson's real criticism.

---

<sup>10</sup> Quine rebukes non-naturalized philosophers for engaging in theoretical "make-believe" by *imposing* conceptual structure onto epistemology, rather than *discovering* structure by attending to the science of psychology (OR, 75).

<sup>11</sup> This is not to say that Quine *accepts* all of Davidson's theories (for instance, unlike Davidson, Quine believes that we can sensibly talk of animals thinking despite lacking language [FTL, 479]), but that he views Davidson's theories, based as they are in physicalism, scientifically (and hence naturalistically) assessable.

Elaborating on his dissatisfaction with Quine, Davidson writes that “[Quine’s] epistemology remains resolutely individualistic...there is no reason in principle why we could not win an understanding of the world on our own” (E, 10), because “[Quine] makes the content of empirical knowledge depend on something that is not shared with others” (CKVP, 291), namely, private stimulations. In Davidson’s view, “this is not a position which can get a grip on the objective character of thought” (E, 10), since the content of our beliefs would then lack a connection with the public world, and any external “control” (RQ, 84). Is it fair to characterize Quine’s epistemology as “individualistic”? Is Davidson right that Quine lacks an account of the “objectivity” of thought?

## **2. Truth, Objectivity, and Solipsism**

My epistemology starts from intersubjectivity, that is, from the experience of sharing objectivity...thinking presupposes intersubjectivity. This will remain an irreconcilable dispute between Quine and me. (Davidson, quoted in Borradori 1994, 53-54)

I have argued that Davidson’s challenge of skepticism about the external world and his doubt about the intelligibility of unconceptualized reasons are ancillary considerations of his critique of naturalism that fail to properly engage Quine. His central objection is rather to the naturalist account of the objectivity of thought. In this section, I shall distinguish Davidson’s position from Ebbs’. In chapter two, I argued that Ebbs motivates his alternative, “participant perspective” on language use by claiming that the naturalist misdescribes the norms grounding our dispute-resolving *linguistic* practices. Yet we saw that it was Ebbs who misdescribes Quine as having an idiolectal conception of language. In contrast, Davidson finds the naturalist account of objectivity *epistemically* wanting. He thinks it inadequate to explain what it means to “get it right” about the world, not because Quine *lacks* an account of getting it right, or of how inquirers putatively justify their knowledge to each other, but rather because Quine cannot

properly explain how inquirers' knowledge counts as *being about* the world at all. This is, I think, the root of Ebbs' misgivings about the naturalist's explanation of our dispute-resolving practices, an account which Davidson likewise finds unappealing. But unlike Ebbs, Davidson has developed an independently motivated, humanist alternative to naturalism, under which our conception of objectivity is dependent upon our intersubjectivity. I will not only show that humanism differs from naturalism in counting solipsism unintelligible, but explain why its privileging of our grasp of the shared concept of truth as a condition for linguistic ability and cognition makes intersubjectivity the *precondition* for any epistemological theory.

To begin clarifying Davidson's position, let us unpack his charge that Quine cannot capture the "objective character" of thought (E, 10). This refers to Frege's claim, discussed in chapter one, that communicators can be said to agree or disagree *only* if they share some "content." In Frege's view, we ought to distinguish the resolution-demanding contradictory attitudes that characterize scientific disputes from the merely differing opinions that underlie disputes about taste, and recognize that it is only in scientific, truth-apt cases that disputants count as agreeing or disagreeing about something. Frege concludes that although each inquirer's *thinking* is subjective, the thoughts *that* each thinks cannot have a purely subjective character.

Neither Quine nor Davidson—nor, I have argued, Frege—posits a category of Platonic, *ontologically* objective entities to satisfy the demand that thoughts not be purely subjective. (On the deflationary reading I have pursued, Frege rather conceives of the "objectivity" of our thoughts in terms of the features that distinguish our practice of inquiry from other practices in which we engage.) But in arguing that translation is indeterminate, both Quine and Davidson reject Frege's idea that inquirers' utterances have univocal meanings at all, be they Fregean "thoughts" or the more contemporary "propositions." Nevertheless, Davidson endorses Frege's



view that communicators must share something if they are correctly to be called communicators. In contrast, Quine rejects Frege's account of communication. His deflationary, behaviorist ersatz of "agreement" as *constituted* by seeming to enjoy "fluent dialogue and successful negotiation" (quoted in Dreben 2004, 289) with one's interlocutor makes communication very thin. Quinean communicators do not mutually grasp and endorse some "content." Their attributions of "beliefs" to each other are fallible parts of their theories for predicting each other's dispositions to vocally react to stimulation by the world.

Quine thinks that he can adequately explain objectivity, and our beliefs and utterances being "about" the world, within this framework. Taking himself to have learnt the lesson Davidson wished to teach him by eliminating shared stimulus meanings from his theory, Quine appropriates Davidson's triangulation metaphor in the title of a very late paper, "I, You, and It: An Epistemological Triangle." There, Quine argues that although we share no stimulations intersubjectively, evolution ensures that both "You" and "I" subjectively respond to causal influence from similar parts of "It," which allows "communication [between us to proceed] apace" (IYI, 486) in Quine's attenuated, behaviorist sense.<sup>12</sup> To Quine, each individual's utterances have the external control of other competent language users, who will object if the home language is misapplied to the shared world.

But Quine here misappropriates Davidson's metaphor. He has begun describing a *semantic* triangle that purports to shed light on how our utterances come to be meaningfully about the world. Quine explains the meaningfulness of our home language socially, interpersonally. Since our beliefs typically find expression in language, Quine takes his triangle

---

<sup>12</sup> Since Quine here says that our ability to find others' utterances meaningful depends on how we are both causally affected at the "It" vertex of the triangle, Davidson now counts him an ally in semantic externalism. However, Quine would object to the intensionalist ring of Davidson's preferred formulation of this doctrine, that our words have the semantic "content" which they do in virtue of causal chains stemming from shared, distal parts of the world to each of us.

to derivatively show how our beliefs concern the world. But Davidson's triangle is intended to be *epistemological* from the outset. He wants to use it to shed light not merely on how our linguistically expressible beliefs are about the world, but on how we are capable of having beliefs that are about the world at all. Whereas Quine continues to valorize the private, subjective capacity to recognize perceptual similarities as "vital...for all learning, all habit formation, all expectation" (IYI, 486)—in short, the *biological* precondition allowing each individual mind to develop theories—Davidson rules intersubjective communication the precondition for there being individual minds. For Davidson, we cannot start epistemology prior to intersubjectivity.

The different uses that Quine and Davidson make of triangulation also highlight the methodological difference between naturalism and humanism. For Quine, triangulation is a way of emphasizing the social nature of language. It makes explicit the role of other people in each individual's introduction into the practice of language use. Since "teaching [a language] in the nursery" is, at the level of observation sentences, "a matter of perceiving that the subject is perceiving that *p*" (PT, 62-63), our language use depends upon mastery of the idiom "he/she perceives that *p*" which makes ineliminable reference to another person (c.f. RR, 10). Quine thinks he has captured the epistemologically relevant points Davidson wished to make by placing triangulation in the context of Darwin's theory of evolution. Evolution accounts for members of the same species carving up their sensory experience of the environment in parallel ways, and so, explains how individuals within a community are able to master the idiom required for language use. For Quine, triangulation features in a *psychogenetic* story of how the human species, and individual humans, develop the capacity to use language. His theory dovetails with scientific theories of developmental psychology and evolutionary biology. In contrast, Davidson's theory

of the acquisition of the shared concept of truth is not psychogenetic, and is not under scientific constraints. His arguments about triangulation are designed to expose *conceptual* interdependencies between the core concepts of epistemology, which, given the success of mutual human interpretation, show that we have the capacity to know about the world, and that that capacity is dependent upon our ability to interpret one another. Davidson's theory is meant to contribute to our self-understanding by forcing us to acknowledge the role of intersubjectivity.

This fundamental difference between their views is why Davidson calls Quine's epistemology "individualistic" (E, 10). Quine's position remains essentially unchanged from an early paper, where he evocatively writes:

I am a physical object sitting in a physical world. Some of the forces of this physical world impinge on my surface. Light rays strike my retinas; molecules bombard my eardrums and fingertips. I strike back, emanating concentric air-waves...our *knowledge* must depend thus solely on surface irritation and internal conditions. (SLS, 1, my emphasis)

Quine thinks that invoking "surface irritation" (or the "stimulations" he later prefers) correctly depicts the epistemological subject as part of the causal structure of the physical world. He also thinks that individuating an individual's perceptions by her stimulations avoids the complications of specifying "the" cause in the external world of our interlocutor's perception at a particular moment (PTF, 475). But from Davidson's perspective, this last restriction thwarts explaining our interpretive attempts to grasp what our interlocutor's—and also our own—beliefs are about.

As we saw in chapter three, Davidson deploys the epistemological triangle to make vivid inquirers' grasp of the contrast between the paired concepts of belief and truth. In late papers where he tries to explain his continued problem with Quine's view, Davidson objects to the limited status which Quine awards the concept of truth. He asks Quine to clarify his "slightly mysterious" (MTE, 57)—and potentially belittling—description of truth as "immanent" (TTPT, 22). Quine responds that he is not claiming truth is *determined* by an individual's current theory

(WWD, 77).<sup>13</sup> Rather, he is emphasizing that cosmic exile is a myth. Truth is judged from within one's current theory, as a part of that theory. The disquotational paradigm for truth ("s" is true if and only if s) uniquely determines the satisfaction conditions for a truth predicate over the sentences with which an individual is competent. It follows that a speaker understands the ascription of truth to a sentence just in case she understands it: her ability to correctly use her truth predicate is secured by her linguistic competency. Quine suggests a genetic explanation for the truth predicate on the basis of its usefulness in communication.<sup>14</sup> He also grants that truth "doctrinally" transcends our theory: we do not say a discredited claim *was* true but *became* false, but that we *thought* it true though it *never was* (77).

Despite all this, since Quine focuses on disquotation and truth's transparent application to one's own sentences, Davidson thinks him blind to the real significance of the concept of truth: its foundational epistemic role in applying to others' utterances. On Quine's account, extending one's truth predicate to someone else's speech is merely a part of developing a theory of translation for her, and choosing a manual that accords with the disquotational paradigm is a powerful heuristic device—and so a significant desideratum—for convenient translation. Davidson thinks that this inverts matters. His transcendental argument involves humanist analyses of "truth" and "belief" that expose sharing the concept of truth to be a condition upon interpretation. Interpretive T-theories may or may not lead to manuals that are neatly disquotational, but they *all* attribute the concept of truth to the person interpreted. The reason is that possessing the concept of truth is necessary for the subject of interpretation (be it another

---

<sup>13</sup> Taking naturalized epistemology as his inspiration, Lars Bergström develops an "empiricist" theory of truth along essentially these lines, which Quine finds "appealing" but "disconcerting," and explicitly rejects (IYI, 489).

<sup>14</sup> We can use a truth predicate to explicitly mention sentences in our utterances ("‘Grass is green’ is true if..."), and thereby avoid use/mention mistakes by signalling to our interlocutor that we are intending to convey what we take the sentence to mean (WO, 272). A truth predicate also allows us to endorse another's assertion without a lengthy repetition (e.g. "That's true"), and to quickly defer to authorities (e.g. "Everything Einstein said is true").

person or oneself) to be understandable as having beliefs about the world, and, according to Davidson, we can only acquire the concept of truth through interpreting others.

Will this argument about the need for epistemology to begin with the shared concept of truth compel Quine? It is helpful to contrast it with the new skeptical objection that Stroud levels against naturalized epistemology. Stroud thinks that, under naturalism, *because* each inquirer cannot check her beliefs concerning the world against an independent standard, no inquirer can justify holding her beliefs true (1984, 244). As we saw in chapter two, Quine can respond to Stroud by leaning on the fallibility of science. Our Humean predicament means we lack certainty about the world, but scientific realism is nevertheless the best theory we have according to the scientific tribunal, which is, in turn, the best standard we have for “justifying” our theories. In contrast, Davidson goes beyond Stroud to insist that the independent standard needed for an individual’s beliefs is knowledge of another’s mind, for it is *only* by interpreting another’s mind that one can be said to have beliefs about the world oneself. So, Davidson’s humanist criticism of naturalized epistemology is that, *because* each inquirer cannot check her “beliefs” concerning the world against the independent standard garnered by another mind, they cannot be counted *beliefs* at all. How might Quine respond?

Quine’s view can easily absorb Davidson’s claim that linguistic animals (and humans in particular) appreciate the truth-belief contrast as they acquire language.<sup>15</sup> But the force behind Davidson’s humanist epistemology is his claims that the *only* intelligible genesis of a mind is its acknowledgement of the truth-belief contrast, which can *only* occur through communication with

---

<sup>15</sup> Indeed, Quine’s appeal to empathy demonstrates that he recognizes the need for a subjective-objective contrast in language acquisition (PT, 42). In Quine’s view, a mother teaching her daughter language must empathetically project herself into her daughter’s position and imagine what aspect of their shared environment she is attending to in order to appropriately reward or admonish her intermittent utterances.

another mind. In an interview with Giovanna Borradori in 1990, Davidson encapsulates his position:

The empiricists have it exactly backwards, because they think that first one knows what's in his own mind, then, with luck, he finds out what is in the outside world, and, with even more luck, he finds out what is in somebody else's mind. I think differently. First we find out what is in somebody else's mind, and by then we have got all the rest. Of course, I really think that it all comes at the same time. (Davidson, quoted in Borradori 1994, 50)

Quine can see no way to establish Davidson's claims naturalistically. Davidson's report that he cannot imagine how the truth-belief contrast *could* arise except via communication is just an introspective claim, scientifically unfit to establish a *standard* for "mindedness." Quine thus appeals to his naturalist methodology to refuse Davidson's distinctively humanist epistemological claim.

But this cuts both ways. The emphasis Davidson lays on our concepts belonging to us, and his resulting inability to conceive of how beings could develop what humans call "minds" with "beliefs" about the external world *except* through linguistic interpretation, undermines the import of Quine's naturalist theory of reification, based as it is on an *individual's* sensory stimulations. So despite being unable to undermine naturalism on Quine's terms, Davidson's objections are not empty. His humanist epistemology captures a thought absent from naturalist epistemology, namely, that being *answerable* to each other about how we believe the world to be is a condition for our beliefs to count as being about the world at all, as opposed to our private, subjective experience.

One way of appreciating this difference is to compare Quine's and Davidson's respective rejections of solipsism. Quine appeals to his deflationary notion of fluency—the lack of conversational breakdown—as the sole criterion of successful translation. There are no semantic constraints upon the recognition of fluent discourse in others; fluency is behaviorally discernible even by those not fluent in the language. To promote fluency in the initial stages of translation,

the translator must suppress the difference between how things are and the way she believes them to be: she reckons her interlocutor's sentences true just in case (in her current manual) they translate sentences *she* calls true. This entails that the corrective epistemic standard provided by other inquirers whom she translates as "disagreeing" with her only holds for as long as she stands by her current manual. In the face of sustained "disagreement," she may develop a new manual rather than seriously question the veracity of her own beliefs.

Exactly the same principle applies to our home language. What favors our (typically unthinking) use of homophonic translation is its tendency to result in fluent discourse with other home language speakers. Since our semantic theory at home is a fallible part of our web of beliefs, the "control" we offer for each other's theories about the world is a standard that can, in principle, be refused. I am not suggesting that Quine thinks that the naturalist *should* refuse this standard: far from it, given the fluent conversations she generally enjoys with others. The point is rather that she *may* intelligibly refuse it. Even if *every* member of her speech community were to object to one of her beliefs, the stubborn naturalist may choose to explain the "disagreement" by supposing that the fault lies in her semantic theory, her translation manual, rather than her own beliefs about the world. She will then have to justify her hypothesis at the scientific tribunal.

So, Quine finds skepticism about other minds to be a *constructive* skepticism that is naturalistically answerable: the possibility that I am the sole inquirer, and that everyone I translate as talking about the world are mere automata (whose regular noises I have mistaken for utterances intended by other minds) is conceivable, but ultimately falls to Darwin. Positing my evolution with fellow inquirers yields the best explanation I (currently) have for my own ability to inquire. Notice that in contemplating this constructive skeptical hypothesis, Quine, ever the

behaviorist, would still think he shares a home language with the automata. Language is a social art; utterances are only meaningful in virtue of the regularities in one's verbal behavior in coordination with that of one's community. If Quine's is the sole mind, he thinks it nevertheless in a shared home language that he acquiesces to questions about ontology (or, better, noises he takes to be questions). The norms underwriting the dispute-resolving behavior in which he engages with the automata—*contra* Ebbs—is explicable by reference to the scientific standards he takes them all to share.

In contrast, Davidson makes imputation of the grasp of the concept of truth the goal of interpretation. The interpreter is not simply projecting *her* application of the truth predicate onto her interlocutor's utterances, but attributing to her interlocutor recognition of an independent standard transcending them both. (Recall from chapter three that this was the force of the principle of charity for Davidson, rather than it being merely convenient to find the natives agreeing with us.) If this attribution is justified, she is *obliged* to treat her interlocutor's beliefs seriously as a possible corrective to her own. This is why Davidson holds disagreement over words to be cause for mutual reinterpretation about the world. Communication is a constant struggle to interpret, and to make ourselves interpretable to, our interlocutor (NDE, 102). Davidson concurs with Ebbs that the onus of justification borne by the individual who judges against the group is not, as Quine thinks, merely a matter of us all being subject to the scientific tribunal. Instead, it is a function of (so we might say) our metaphysical epistemic situation: we are inquirers who count as seeking objective truth only because we recognize that none of us have a privileged grasp of how the world really is.

Davidson rules the solipsistic threat (i.e., that *every* interpretation manual I use is mistaken since I am the sole inquirer) *radical*, not constructive. He thinks solipsism based on a



semantic error. Solipsism cannot be formulated without ascribing a mind to oneself, which requires ascribing the truth-belief contrast to oneself on the basis of the epistemological triangle, which in turn demands ascribing beliefs to another mind. While solipsism is naturalistically answerable, it is humanistically unintelligible.

It is worth emphasizing that although I have revealed Davidson's opposition to Quine to be principled, his humanist insistence that other inquirers constitute a robust epistemological standard to which each individual theorist is responsible is not a knockdown argument against Quine. Quine can simply refuse to accept that this standard is as robust as Davidson claims and embrace the "individualistic" inquiry Davidson abhors as the epistemic lot of the naturalist. I ended the previous section by noting that, from Quine's naturalist perspective, Davidson's focus is the "far end" of the causal chain that extends from the world through our sense organs to eventually bloom into our scientific theories about the world. To Quine, Davidson's radical interpretation is an interesting (and naturalizable) account of how linguistic animals develop *theories* of each other as minded. In contrast, Quine's empiricism is uninteresting from Davidson's humanist perspective because it describes stages prior to socialization. He thinks that no organism can count as acquiring "knowledge" of the objective world on the basis of its private sensory stimulations, so Quine's naturalized "epistemology" is merely a metaphorical mentalistic extension of biology. Davidson believes that we will only understand ourselves, and our position within our world, once we leave Cartesianism behind. Our capacity to think and talk—our status as minded communicators—depends on our interpretation of each other.

To further sharpen the difference between Quine and Davidson, I should like to end this section by returning to the thought experiment with which I began this dissertation. Suppose that

we were to encounter a roughly human-looking alien species who appear to be communicating.<sup>16</sup>

An intrepid linguist starts trying to decode their language. A philosopher friend suggests she look to Quine's and Davidson's radical linguist experiments for guidance. Now, the interest of this case surfaces when we suppose the linguist gets into serious difficulties. Try as she might, she cannot devise an entering wedge into the alien tongue. Whatever initial hypotheses she makes about its structure, she quickly encounters counterexamples. She begins to suspect that the alien language is uninterpretable.

To Davidson, this suspicion reveals a semantic error. He thinks "an uninterpretable language," like "a radically different conceptual scheme," is a contradiction in terms. *Either* the aliens lack language and are unminded *or* they are using a language that we are currently experiencing difficulty interpreting; there is no third possibility. The linguist's error is tempting because we cannot speak every language of which we are aware, and it seems only a small step to allow the possibility of languages so different from our own that they exceed our ability to learn them. But Davidson insists on inverting this dependence: something being a language is *predicated* on our ability to interpret it as meaningful.

Davidson criticizes Quine's discussion of this case:

Quine says we might see that members of some group, from outer space, perhaps, are fluently conversing, though we could find no way to map our entities onto parts of their sentences. But how would we identify what we were witnessing as conversation? (RQ, 81-81)

Despite Davidson's dissatisfaction, Quine's evaluation of the case seems extremely similar to his own. Quine would stridently encourage the disheartened linguist to consider her hypotheses

---

<sup>16</sup> Discussing a similar case, Ludwig Wittgenstein suggests we may have evidence that the aliens are communicating if, when we gag them, "their actions fall into confusion" (PI, §207). There are a surprising number of weird and wonderful aliens that philosophers have imagined encountering to test their intuitions about language ascription (see especially Neil Tennant [1999] and Tomáš Marvan [2003]). Nevertheless, since my aim here is to explain how Quine and Davidson differ in evaluating these sorts of cases rather than to defend a particular claim in the philosophy of language, I shall restrict my focus as far as possible to this relatively pedestrian "human-like" alien species.

scientifically. Having chided her for suspecting that the aliens speak an *uninterpretable* language (since it is possible that future linguists will succeed where she has failed), he would defend the more conservative hypothesis that the alien language currently resists interpretation.<sup>17</sup>

Although neither Davidson nor Quine tolerates the supposition of an “uninterpretable language,” the tones of their rejections importantly differ. Echoing Frege’s dismissal of resolute logical aliens as mad, the humanist’s ascription of mindedness to her interlocutor is inseparable from recognizing her interlocutor as using a language that cleaves to logical norms, one that she can employ to refer to objects in the shared world. While Frege wrote metaphorically that the laws of logic “unfolded” the content of the word “true” (1979, 3), Davidson’s theory is an account of *why* logical aliens must be excluded from what we can call having cognition (and speaking a language): if a group of beings are uninterpretable, in the sense that first-order logic *cannot* be read into their language, then we cannot construct a truth theory for them. It follows that they lack the truth-belief contrast fundamental to our practice of inquiry, and that they have no capacity to inquire about the objective world.

In contrast, from Quine’s naturalist perspective, the import of first-order logic for the philosophy of language and mind is exhausted in recognizing that fixing the logical structure of the alien language is (currently) crucial to radical translation. More generally, the significance of logic to naturalized philosophy is in characterizing the entailment relation demanded by our current scientific theories. Yet, because our scientific theories are fallible and revisable, so too is the logic at their core. If we were to develop scientific theories better able to explain our world

---

<sup>17</sup> Quine may add that the linguist ought to consider alternative hypotheses. Perhaps the alien noises merely function as warning cries, like the cries of seagulls. We might, that is, be able to locate alien “observation sentences” and translate them after a fashion (perhaps they always cry “gavagai” when near a predator), but, because we cannot find any logical structure in their noises, be disinclined to call their cries “language” (unless we wish to say seagulls use language). The details of the case will determine which hypothesis is preferable. To take just one example, the length of observation will be a relevant parameter. If the linguist is venting her frustration after working for a few hours, we may fault her; if she has been working for a few years, we may question the aliens’ capacity for language.

than our current theories, yet which could not be regimented into our current canonical logic, we ought to excise both our current theories and our current logic from our web of belief. Quine would describe this as a paradigm shift, an evolution in our conceptual scheme. Since Quine thinks scientific development might require us to extract even central beliefs from our web, he recognizes no firm link between first-order logic and mindedness.

Unlike Davidson, Quine thus accepts the possibility that aliens might use a language that is uninterpretable (in the sense that we cannot project our logic onto their speech). To Davidson, Quine's naturalistic openness stretches the application of our words to their breaking point. Beings whose "language" lacks the quantificational structure necessary for us to develop an interpretive T-theory—the only way we know of to create a truth theory for a language—are not beings who we cannot understand, but beings who *cannot be understood*, in virtue of lacking what we call "beliefs."

This case indicates a Davidsonian argument against radically different logics that parallels his argument against radically different conceptual schemes. An epistemologist who, like Quine, supposes that we develop theories about the world *using* some logic cannot answer a new sort of skeptic, for, on this view, logical aliens may be developing theories about the world using a "logical" principle that *we* regard unfounded. But having recognized that from *their* perspective they are doing nothing wrong, we can wonder whether we are in a similar predicament, developing theories about reality in a way that the aliens think unfounded. We can go on to wonder whether our logic is the *right* logic, and so, come to question whether it even makes sense to say that an inquirer *has* a right logic.<sup>18</sup>

---

<sup>18</sup> It should be clear that Quine would not be troubled by this new skeptical argument. To the naturalist who appreciates that all science is fallible, wondering "whether our logic is the *right* logic" upon encountering such aliens (where "rightness" is just adequacy for future theorizing) is the appropriate attitude to have.

Davidson would have us reject the dualism upon which this skeptic relies. Just as there are no schemes by which humans organize empirical content, so, too, there are no logics by which inquirers theorize about that content. Humans reason logically about the world with which they are in unmediated contact, since to reason just *is* to obey the prescriptions of our logical laws.

### **3. Resolving Philosophical Disagreement**

I have extracted the following distinct conceptions of logic's relationship to cognition from Quine and Davidson:

**NAT:**        **A person must exhibit speech behavior that can be translated as reasoning in conformity with the laws of our canonical logic if we are to engage in inquiry together.**

**HUM:**        **Attributing the concept of truth to another mind using quantificational logic to structure the other mind's language is a necessary condition for being minded.**

Quine's naturalist conception denies the necessary connection between cognitive ability and facility with our logic that Davidson's humanist conception affirms. The competing conceptions are each grounded in a nuanced approach to philosophy that entails claims in philosophical methodology, epistemology, and the philosophy of language and mind. Which, if either, approach ought we to adopt?

The disagreement between Quine and Davidson is of particular interest in the context of recent attention to the epistemology of disagreement.<sup>19</sup> When equally competent peers with access to identical evidence disagree, to what extent should their confidence in the truth of their own positions be shaken? Is each under an epistemic obligation to learn about alternatives to

---

<sup>19</sup> See Thomas Kelly (2005), David Christensen (2007), and the contributions to Richard Feldman and Ted Warfield (2010).

their view? Most of the current literature on disagreement focuses on moral and religious cases. In this section, I hope to use Quine's and Davidson's dispute to draw some lessons about the case of philosophical disagreement.

The hopeful thought that I want to pursue is that although Quine and Davidson have important and philosophically revealing disagreements, the projects in which they engage do not conflict; they only differ. But this hope looks liable to be dashed in view of their various conflicting claims. In the above sections, we have seen that both are committed to calling their own projects "epistemological," and denying the other that label. Both propose methodological constraints upon properly "philosophical" inquiry, and neither develops theories that meet the other's requirements. Nor are their views terminological variants: the differences between their conceptions of logic and cognition lead them to disagree about the limits of intelligibility, resulting in a variety of low-level conflicts about the meaningful application of particular concepts. For instance, Davidson cannot accept Quine's receptivity to a finite natural language *counting* as a "language," and Quine cannot accept Davidson's restriction that attributing grasp of "truth" to one's interlocutor is "necessary" for communication. Some work is required to explain why these conflicting claims do not entail that Quine's and Davidson's epistemological projects conflict.

Firstly, it is worth emphasizing that the disagreement between Quine and Davidson is *intractable*. Some disputes can be resolved by appealing to a standard on which disputants agree. For example, a dispute about the number of beans in a jar can typically be resolved by tipping out the beans and counting them. Some philosophical disputes are similarly resolvable. For instance, purported ontological proofs of God's existence can be scrutinized for soundness and validity and found wanting. But the dispute between naturalism and humanism is not so

clear-cut. Each approach counts different considerations persuasive, different cases intelligible, different methods properly philosophical, and characterize not only the goal of rational inquiry, but our *grasp* of that goal, differently. Arbitrating between them using the standards of evaluation which they share is inconclusive; no demonstrative argument favors one account while undermining the other. But arbitrating using a standard of evaluation that only one approach endorses begs the question against the other.

We saw an example of this in section two. Quine might object that Davidson's argument for the necessarily interpersonal origin of the truth/belief distinction is introspective and unscientific, and Davidson may count Quine's conception of objectivity inadequate to explain our status as subjective inquirers who share a world. Neither objection is persuasive from the other's perspective. The naturalist can defend her austere epistemology at the scientific tribunal, and the humanist can defend her social epistemology by appealing to her transcendental argument upon the conditions of interpretation. This parallelism results in a philosophical standoff, and makes any naturalist or humanist who thinks their own perspective *true* and the other *false* look like an objectionable philosophical egoist.

Writing about Quine, P.F. Strawson suggests:

[T]he choice between [Quine's naturalism and ordinary language philosophy] is ultimately, perhaps, a matter of individual temperament; and if I have made my own preference clear, it is no more than that—my own preference. It has been said that the best conceptual scheme, the best system of ideas is the one that gets us around best...For one content to lead his life—at least his intellectual life—in the rarefied atmosphere of science, the choice...will go one way. For one content to lead his intellectual life in the muddier atmosphere of the more mundane...it will go the other. (1990, 318, original emphasis)

If Quine's preference for naturalism over humanism—or Strawson's ordinary language philosophy—is ultimately a matter of his temperament, we have reason to think that viewing his disagreement with Davidson as a *conflict* between their views, at most one of which can be *true*, is wrongheaded. Their disagreement is intractable because it is grounded in differences between

their preferred form of inquiry, preferences which brook no argument. The situation recalls Carnap's metaphilosophy. In *The Logical Syntax of Language*, Carnap argues that some philosophical disputes can only be resolved indirectly. Inquirers who share a temperament will pragmatically agree upon an argumentative framework and will mutually benefit from engaging each other. But to engage an inquirer who disagrees about the framework in which arguments are to be evaluated on anything but pragmatic grounds is futile. Carnap recommends that we instead *tolerate* an inquirer whose argumentative framework (and temperament) we do not share. Once competing projects have developed in tandem, the theories generated under each perspective can be compared. A pragmatic adjudication between temperaments may then be possible.<sup>20</sup>

The present disagreement cannot be captured using Carnap's syntactic method. He argues that inquirers must agree on which truths they consider "analytic" before they can debate, yet Quine and Davidson both reject the analytic/synthetic distinction. Nevertheless, if the dispute between naturalism and humanism comes down to opposed temperaments, we might think that these disputants warrant a similarly tolerant attitude. Carnap himself sought philosophical "progress," hoping to shift argumentative focus away from (what he saw as) fruitless metaphilosophical discussions to clearly demarcated philosophical programs. Should we likewise postpone judgment about whether naturalism or humanism are "true," and pragmatically decide in the future which approach has allowed for the best theories?

I think that this would be a mistake. One reason is that delaying often promotes sideling, and this dissertation has demonstrated that a grasp of naturalism and humanism (and

---

<sup>20</sup> This is how Carnap suggests we ought to treat the disagreement between classical and constructive mathematicians. Philosophers should choose formal language I or II depending on their temperament, tolerate those who make a different choice, and get on with the business of analyzing scientific language. At some future time, philosophers could reflect on which logic had allowed for the best explanation of science (2002, 164-165; 332).



constitutivism) are necessary to properly understand the arguments developed by Quine and Davidson (and Frege). We are thus in danger of misunderstanding our recent history if we elect to ignore this dispute. But a second danger is that, in this case, Carnapian tolerance encourages the fragmentation of philosophy into distinct, incommensurable schools. Although Quine and Davidson focus on particular problems in logic, epistemology, and the philosophy of language and mind, the *methodological* consequences of their views about the correct way to pursue philosophical inquiry mean that their disciples can apply naturalism or humanism to various problems across philosophy. Unlike the disputants that interest Carnap, naturalists and humanists do not share standards of success, and so, there is no reason to think that they will agree in the future about which approach's theories are better. Tolerating another's naturalism or humanism at a distance—thinking her temperament beyond the reach of rational argument—inhibits dialogue between philosophers of different stripes.

A similar problem affects a metaphilosophical position that seeks to justify holding one's own philosophical approach *correct* or *true* even in the face of competitors. In a moral case that parallels the philosophical one I am discussing, Kieran Setiya (forthcoming) argues that “epistemic localism” is necessary to avoid the twin pitfalls that come with an evidential moral epistemology: parochial egoism (thinking one's own moral theory correct) and unsatisfactory skepticism (thinking no moral theory is correct). The epistemic localist denies that “the fundamental standards of epistemic justification are topic neutral” (19) and takes moral education to involve learning the standards of epistemic justification in morality. This entitles her to hold her own moral theory fallibly correct, based on the perceived efficacy of her moral education, and to justly rebuke “moral monsters” (19) with different, purportedly moral views. Setiya suggests that epistemic localism is necessary to overcome egoism and skepticism

wherever radical disagreement occurs (n27). So, might philosophical education likewise involve learning the standards of epistemic justification in philosophy, entitling those taught naturalism and those taught humanism to each hold their approaches fallibly correct? The danger with such epistemic localism is that it gives inquirers no incentive to learn about other approaches and the insights those approaches might yield when applied to particular problems. It is a desideratum of one's metaphilosophical theory that some such incentive is in place, so as to avoid licensing parochial bias of one's own education while nevertheless maintaining a tolerant attitude.

To clarify the position I propose, let us consider the metaphilosophical attitudes that Quine and Davidson would have to their own disagreement. I think that we can reconstruct such attitudes from their late disagreement about empirically equivalent theories (PT, 95-102).<sup>21</sup> Call two theories *empirically equivalent* if and only if whatever observation counts for (against) one theory counts for (against) the other. Call a theory a *global system* if it is a theory about our whole world that claims to fit all of our possible observations. Now suppose two scientists, Jerome and Janine, discover differences in their empirically equivalent global systems.<sup>22</sup> Their vocabularies—ranging from observational through to theoretical—are (phonologically speaking) identical.<sup>23</sup> But although they agree upon all observation sentences, such as “mailboxes are blue,” each denies theoretical sentences that the other affirms.

In one case, once their theories are presented systematically, structural similarities between them emerge—for instance, perhaps Jerome consistently uses “proton” when Janine

---

<sup>21</sup> Beyond mentioning their opposed views, Quine does not go into very much detail about the cases that he, Davidson, Dagfinn Føllesdal, and Burton Dreben discussed in their closed conference at Stanford in 1986. Since, in contrast with their interest in the scientific principles for theory construction, my interest in their dispute is metaphilosophical, I shall make certain stipulations to make the case that interests me more precise.

<sup>22</sup> Let us stipulate that, on the basis of inductive reflection on their past conversations, Jerome and Janine are confident that their theories are empirically equivalent.

<sup>23</sup> Although some words are clearly observational (“red,” “round”) and some are clearly theoretical—meant to explain unobserved reality (“quark,” “dark matter”)—Quine and Davidson both think that observational and theoretical vocabulary differ only in degree and not in kind.

uses “neutron” and vice versa—which means that they can treat their theories as terminological variants. But in a second case, Jerome and Janine find no way to map their theoretical vocabularies onto each other. Since they agree about observational sentences, Janine decides to teach Jerome her theory of natural science from the ground up. In time, Jerome is able to reliably predict how Janine uses her theoretical sentences and to do science with her. Nevertheless, because their competing theories are phrased using phonologically identical vocabulary, Jerome finds himself denying sentences when using Janine’s theory that previously he had affirmed. Jerome judges the two theories to be on a par according to the principles of theoretical construction: they are equally simple, conservative, etc. (although his own is a good deal more familiar). What should Jerome’s attitude be toward the *truth* of his two theories?

Jerome and Janine’s situation bears an important resemblance to Quine and Davidson’s. Quine and Davidson agree upon observation sentences—and, unlike Jerome and Janine, a good deal of empirical science—but *not* about philosophy. Their different naturalist and humanist views are phrased in the same vocabulary (each makes claims about “truth,” “epistemology,” etc.), are not terminological variants, and entail that there are sentences that one affirms and the other denies (such as **HUM**). Were Quine and Davidson to learn how to speak each other’s philosophy, what should their attitude be toward the *truth* of humanism and naturalism?

Although Quine fluctuated on the issue of whether or not to hold empirically equivalent theories true,<sup>24</sup> he came to favor a “sectarian” view: Jerome should hold his theory true and Janine’s false, but feel free to swap and use Janine’s theory (holding his own theory false) if the new perspective he gains in doing so is helpful. Davidson disagreed. On his “ecumenical” alternative, Jerome should use subscripts to eliminate the formal inconsistency that seemed to threaten his own utterances (since he affirms sentences when using his theory that are

---

<sup>24</sup> See RIT, 178-183; EC, 28-29; PT, 100.

phonologically identical to those he denies when using Janine's) and hold both theories true within a larger, inclusive language.

Quine came to favor sectarianism over ecumenism because naturalism trumps empiricism (PT, 100). As an empiricist, Quine finds ecumenism attractive since both theories are empirically adequate, and so Jerome and Janine are equally warranted in believing them. However, naturalism requires Jerome to follow the standards of the scientific tribunal. Simplicity favors choosing *either* his own theory *or* Janine's, since using subscripts as Davidson suggests to create a combined theory that incorporates both sets of theoretical vocabulary is ontologically extravagant. Jerome would have to admit, for instance, that  $\text{protons}_{\text{JEROME}}$  *and*  $\text{protons}_{\text{JANINE}}$  exist, yet just one theory was sufficient to explain the observable data. Moreover, the familiarity of his own theory points in its favor. Having learned Janine's theory, Jerome can always adopt it for a time if he needs to do theoretical research with her—and can certainly appeal to it if doing so is helpful. In other cases, since they do not disagree about non-theoretical sentences, he is entitled to hold his own global theory true and hers false. When phrased in canonical logic, Janine's theory wrongly endorses existential claims that Jerome's theory denies.

Quine is more tentative about his sectarian attitude to empirically equivalent theories than he is to allowing acquiescence in the home language to questions of ontology. Unlike the proxy functions that make cases of deviant ontologies precise, imagining structurally distinct, equally simple and conservative, empirically equivalent global theories is far-fetched. If we were really to be confronted with two such theories, the attitude that we should take to their truth would depend upon the details of the case.<sup>25</sup> Nevertheless, it is clear why naturalism leads Quine in the direction of sectarianism. Why does Davidson prefer ecumenism?

---

<sup>25</sup> Alexander George makes a similar use of Quine's discussion of empirically equivalent theories to question Quine's view of his disagreement with Carnap (2000). George quotes Quine: "the cosmic question whether to call

As Davidson views matters, Jerome and Janine have attributed grasp of the shared concept of truth to each other by mutually interpreting observation sentences. But if Jerome adopts a sectarian attitude, he effectively treats Janine as massively mistaken in her theoretical beliefs about the world. He adopts a *proprietary* attitude to the concept of truth.

A proprietary attitude to truth is legitimate if one can justify thinking that one has better access to truths than one's interlocutor. For instance, an interlocutor with a theological global theory that has weak predictive power is justly refused. But Janine's theory of unobserved reality is not wantonly extravagant or mysterious; it has exactly the same predictive power as Jerome's own. Given this, Davidson's humanism denies that Jerome is entitled to hold himself as having privileged access to truths. The objectivity of truth is explained in terms of our shared intersubjectivity, and acknowledging our metaphysical epistemic situation, as I put it above, requires recognizing that truth is the shared goal of all inquirers to which *no* inquirer has privileged access. Once we ascribe grasp of the concept of truth to an interlocutor, we are obliged to hold her beliefs true as far as we possibly can.<sup>26</sup>

We find a second reason for ecumenism in the formal constraints of interpretation. Davidson's interpreter is not trying, like Quine's translator, to further shared projects with her interlocutor, but to acknowledge her interlocutor as a rational being. Suppose an interpreter tries to interpret a sectarian student of Jerome and Janine—Justin—who has learnt both of their global

---

two such world systems true [simmers] down, bathetically, to a question of words" (PT, 100-101), and concludes that Quine "opts, as it were, for an ecumenical approach to the dispute between the sectarian and the ecumenist" (21). But the quote does not sustain George's reading. Quine takes the *abstract* dispute between the sectarian and the ecumenist to be a question of words, given that a *real* decision about what to do with two given empirically equivalent global systems would depend upon the details of the case (such as how many of our interlocutors use the alternate theory, for example). Nevertheless, I think that he sees a genuine difference between sectarianism and ecumenism applied to particular cases, and makes his own preference for sectarianism (given naturalism) clear.

<sup>26</sup> Quine acknowledges "the satisfaction of conferring the cachet of truth evenhandedly" but insists that since "the sectarian...can still be evenhanded with the cachet of warrantedness" the gains of ecumenism are "not evident" (PT, 100). His thought is that Jerome can think Janine warranted in holding her theory true while nevertheless judging it wholly false. But from Davidson's view, Jerome badly misunderstands his epistemic situation if he fails to confer the cachet of truth evenhandedly between every inquirer who he interprets as grasping the shared concept of truth.

theories, finds them equally familiar, and who has decided to switch between them when useful. To avoid uncharitably concluding that Justin holds true contradictions (since he affirms sentences using one theory that he denies when using the other) the interpreter isolates the two theories Justin is using and gives truth theories for them separately. Abiding by the principle of charity, she will count both of Justin's theories mostly true. Hence, she will interpret Justin ecumenically, holding both theories true in the terms of her language. Since this interpretive policy will go as well for self-interpretation as for the interpretation of others, we ought to be ecumenical if we learn two empirically equivalent global theories ourselves.

In terms of their own metaphilosophical disagreement, Quine's naturalism will lead him to be sectarian about their dispute, while Davidson's humanism will lead him to be ecumenical about their dispute. Quine will count himself entitled to hold naturalism true and humanism false, while swapping to a humanist perspective if he finds doing so helpful. Davidson will hold both naturalism and humanism true within a language that encompasses them both.

Quine's sectarian metaphilosophical attitude might seem to avoid the problem of fragmentation that I mentioned earlier. Naturalists have some incentive to learn about humanism, because it might offer them a new, helpful perspective on certain problems. But, on the other hand, the problems that one thinks one must face are intertwined with one's perspective. Quine, notably, was quite at peace with naturalism, happy, as he wrote in his last paper, to "confess" his "confirmed extensionalism" (CCE, 498). He found little of value in the problems occupying non-naturalized first philosophers. So the incentive for naturalists to learn about humanism under sectarianism does not amount to a *requirement* to adopt humanism for certain investigations, nor to consider the problems that humanist philosophers face. Since

naturalist methodology can be applied across philosophy, the naturalist who is convinced that she has little to learn from humanism is still licensed to work in isolation.

In contrast, Davidson's ecumenical metaphilosophical attitude forces us to acknowledge that other people have a grasp of truth, and that what at first looks to us like others' eccentricities may be important insights. Choosing to restrict ourselves to the approach we find familiar to determine which problems to consider, and how those problems are best considered, becomes parochial. The onus is placed upon us to familiarize ourselves with other schools of thought and to determine, by temporarily adopting their approach, which problems demand our attention and whether a different perspective will help us with particular problems. The strength of this form of tolerance lies in reminding us to look at philosophical problems in more than one way, determining the strengths and weaknesses of each method for approaching them as we do so. This avoids the fragmentation of philosophy into incommensurable schools by forcing us to adopt that perspective which is demanded by our current inquiry.<sup>27</sup>

Although I cannot explore the connections in detail here, the form of tolerance I am defending is similar to John McDowell's quietism (2009), as well as recent interpretations of Wittgenstein's late methodology.<sup>28</sup> These connections expose the difference between my position and Gary Ebbs'. In chapter two, while praising Gary Ebbs' attention to metaphilosophy in thinking about Quine, I criticized his argument that naturalism was *incapable* of capturing our dispute-resolving practices. Although Ebbs begins his book with the Wittgensteinian rhetoric

---

<sup>27</sup> Although I am defending the metaphilosophical ecumenism that Davidson's humanism promotes, I think that there are numerous philosophical problems that we ought not approach as humanists. For instance, to develop a meaningful environmental ethic I think we must see ourselves as *part* of nature rather than in a unique, privileged, and dominant humanist epistemic position. Approaching problems in environmental ethics humanistically is objectionably anthropocentric. In contrast, I think we ought not be naturalists when approaching problems in social philosophy. Since contesting membership in the category of "human" is the core of numerous debates about our obligations to each other, such as abortion, slavery, and marriage equality, approaching social problems naturalistically is objectionably misanthropic.

<sup>28</sup> See particularly Oskari Kuusella (2008) and Marie McGinn (2009).

that he hopes to “loosen the hold” of naturalism (2001, 1) and ends with the recognition that “no-one who has ever been in the grip of Quine’s picture of meaning and assertion will be immediately convinced by the foregoing reasons for resisting it” (309), I think that he fails to live up to his own metaphilosophical ambitions by claiming that his arguments amount to *criticisms* of Quine (4; 92). Similarly, he objects that if McDowell’s criticisms of Crispin Wright on rule-following are intended not as arguments but as attempts to “show us a different way of looking at our linguistic practices,” then they are too “elusive” to be persuasive and fail to explain why “Wright’s picture is unacceptable” (89-90). What Ebbs misses is that Wright’s picture *is* acceptable from Wright’s perspective. The rhetoric of “criticism” and “argument” does not find application at the metaphilosophical level, because the *approach* for arguing about this topic is precisely what is at issue. According to McDowell, all the quietist philosopher *can* do is to try to show a different way of looking at things by elucidating the best (and, prior to infection by a particularly compelling view, already natural) way to think about particular cases (2009). If we find the quietist’s remarks elusive, it is up to us to reinterpret them, rather than to reboot the beguiling rhetoric of argument and criticism.

In my view, disciples of Quine and Davidson should not waste time trying to develop arguments aimed to convince each other that naturalism or humanism is *correct*. Instead, they should learn both naturalism and humanism, examine problem cases by applying the methodology each view commends, and evaluate how features of particular cases demand different approaches. I suspect the same advice applies to other, seemingly intractable cases of philosophical disagreement.



## **Bibliography**

- Barrett, Robert and Roger F. Gibson, eds. *Perspectives on Quine*. Oxford: Blackwell, 1990.
- Bergström, Lars. "Davidson's Objections to Quine's Empiricism." Kotatko et al 2001, 17-36.
- Bhushan, Nalini. "The Possibility of a Radically Different Language." *Philosophical Investigations*, 19, (1996): 237-263.
- Borradori, Giovanna. *The American Philosopher*. Trans. Rosanna Crocitto. Chicago: University of Chicago Press, 1994.
- Carnap, Rudolf. *The Logical Syntax of Language*. Trans. Amethe Smeaton. London: Kegan Paul, Trench, Trubner & Co., trans. of *logische Syntax der Sprache*. Vienna: Julius Springer Press 1934. Illinois: Open Court Press, 2002.
- de Caro, Mario, ed. *Interpretations and Causes: New Perspectives on Donald Davidson's Philosophy*. London: Kluwer Academic Press, 1999.
- Christensen, David. "Epistemology of Disagreement: The Good News." *Philosophical Review*, 116, 2 (2007): 187-217.
- Davidson, Donald. "A Coherence Theory of Truth and Knowledge." (CTTK) *Kant oder Hegel?* Ed. Dieter Heinrich. Stuttgart: Klett-Cotta, 1983. 423-438. Rpt. in Davidson 2006. 225-237.
- \_\_\_\_\_. "Comments on the Karlov Vary Papers." (CKVP) Kotatko et al 2001. 285-308.
- \_\_\_\_\_. "A Nice Derangement of Epitaphs." (NDE) *Philosophical Grounds of Rationality*. Eds. R. Grandy and R. Warner. Oxford: Oxford UP, 1986. Rpt. in Davidson 2006. 251-266.
- \_\_\_\_\_. *The Essential Davidson*. New York: Oxford UP, 2006.
- \_\_\_\_\_. "Externalisms." (E) Kotatko et al 2001. 1-16.
- \_\_\_\_\_. *Inquiries Into Truth and Interpretation*. Oxford: Clarendon Press, 1984.
- \_\_\_\_\_. "Meaning, Truth and Evidence." (MTE) *Perspectives on Quine*. Eds. R. Barrett and R. Gibson. Oxford: Blackwell, 1990. Rpt. in Davidson 2005. 47-62.
- \_\_\_\_\_. "Reply to W.V. Quine." (RQ) Hahn 1999. 80-87.
- \_\_\_\_\_. "On the Very Idea of a Conceptual Scheme." (OVICS) *Proceedings and Addresses of the American Philosophical Association*, 47 (1974): 5-20. Rpt. in Davidson 1984. 183-198.
- \_\_\_\_\_. *Subjective, Intersubjective, Objective*. Oxford: Oxford UP, 2001a.
- \_\_\_\_\_. *Truth, Language and History*. Oxford: Clarendon Press, 2005.
- \_\_\_\_\_ and Jaakko Hintikka eds. *Words and Objections: Essays on the Work of W.V. Quine*. Holland: D. Reidel, 1969.
- Dreben, Burton. "Quine on Quine." Gibson Jr. 2004, 287-293.

- Ebbs, Gary. *Rule-Following and Realism*. Harvard: Harvard UP, 2001.
- Feldman, Richard and Ted Warfield (eds). *Disagreement*. Oxford: Oxford UP, 2010.
- Frege, Gottlob. *Posthumous Writings*. Eds. Hans Hermes, Freidrich Kambartel, and Friedrich Kaulbach. Trans. Peter Long and Roger White. Chicago: University of Chicago Press, 1979.
- Gibson Jr., Roger F., ed. *The Cambridge Companion to Quine*. Cambridge: Cambridge UP, 2004.
- Glock, Hans. *Quine and Davidson on Language, Thought and Reality*. Cambridge: Cambridge UP, 2003.
- Hahn, Lewis, ed. *The Philosophy of Donald Davidson*. Chicago: Open Court, 1999.
- Kelly, Thomas. "The Epistemic Significance of Disagreement." *Oxford Studies in Epistemology*. Eds John Hawthorne and Tamar Gendler. Volume 1. Oxford: Oxford UP, 2005.
- Kotatko, P., P. Pagin and G. Segal eds, *Interpreting Davidson*. Stanford: CSLI, 2001.
- Kuusela, Oskari. *The Struggle Against Dogmatism: Wittgenstein and the Concept of Philosophy*. Massachusetts: Harvard UP, 2008.
- Marvan, Tomáš. "Interpretability, Perceptual Schemes and Triangulation." *Sats – Nordic Journal of Philosophy*, 4, 2, (2003): 93-107.
- McDowell, John. *Mind and World*. Cambridge: Harvard UP, 1994.
- \_\_\_\_\_. "Wittgensteinian 'Quietism'." *Common Knowledge*, 15, 3 (2009): 365-372.
- McGinn, Marie. *Elucidating the Tractatus: Wittgenstein's Early Philosophy of Logic and Language*. Oxford: Oxford UP, 2009.
- Quine, W.V. "Confessions of a Confirmed Extensionalist." (CCE) *Future Pasts*. Eds. Juliet Floyd and Sanford Shieh. Oxford: Oxford UP, 2001, 215-221. Rpt. in Quine 2008a, 498-506.
- \_\_\_\_\_. *Confessions of a Confirmed Extensionalist and Other Essays*. Eds. Dagfinn Føllesdal and Douglas Quine. Cambridge: Harvard UP, 2008a
- \_\_\_\_\_. "Empirical Content." (EC) Quine 1981. 24-30.
- \_\_\_\_\_. "The Flowering of Thought in Language." (FTL) *Thought and Language*. ed. John Preston. Cambridge: Cambridge UP, 1997. Rpt in Quine 2008a, 478-484.
- \_\_\_\_\_. *From Stimulus to Science*. Cambridge: Harvard UP, 1995.
- \_\_\_\_\_. "I, You and It: An Epistemological Triangle." (IYI) *Knowledge, Language and Logic: Questions for Quine*. Eds. Alex Orenstein and Petr Kotatko. Boston: Kluwer Academic Publishers 2000, 1-6. Rpt. in Quine 2008a, 485-492.

- \_\_\_\_\_. "Naturalism; Or, Living Within One's Means." (NLWOM) *Dialectica*, 49 (1995): 251-263. Rpt. in Quine 2004. 275-286.
- \_\_\_\_\_. "On the Very Idea of a Third Dogma." (OVITD) Quine 1981. 38-42.
- \_\_\_\_\_. *Philosophy of Logic*. 2nd ed. Harvard: Harvard UP, 1970.
- \_\_\_\_\_. "Progress on Two Fronts." (PTF) *Journal of Philosophy*, 93, 4, (1996): 159-163. Rpt in Quine 2008a, 473-477.
- \_\_\_\_\_. *Pursuit of Truth*. (PT) Rev ed. Cambridge: Harvard UP, 1992.
- \_\_\_\_\_. *Quine in Dialogue*. Eds. Dagfinn Føllesdal and Douglas Quine. Cambridge: Harvard UP, 2008b.
- \_\_\_\_\_. "On the Reasons for Indeterminacy of Translation." (RIT) *Journal of Philosophy*, 67 (1970): 178-183.
- \_\_\_\_\_. "Replies." (R) Davidson and Hintikka 1969, 292-352.
- \_\_\_\_\_. *The Roots of Reference*. (RR) LaSalle, Illinois: Open Court, 1973.
- \_\_\_\_\_. "Structure and Nature." (SN) *The Journal of Philosophy*, 89, 1, 1992. 1-9.
- \_\_\_\_\_. "The Scope and Language of Science." (SLS) *The Unity of Knowledge*, ed. Lewis Leary. New York: Doubleday, 1955. Rpt in Quine 1966, 215-232.
- \_\_\_\_\_. *Theories and Things*. Cambridge: Belknap Press, 1981.
- \_\_\_\_\_. "Things and Their Place in Theories." (TTPT) Quine, 1981. 1-23.
- \_\_\_\_\_. "Three Indeterminacies." (TI) Barrett and Gibson 1990. 1-16.
- \_\_\_\_\_. *The Ways of Paradox and other essays*. New York: Random House, 1966.
- \_\_\_\_\_. "Where Do We Disagree?" (WWD) Hahn, 1999. 73-79.
- \_\_\_\_\_. *Word and Object*. (WO) Harvard: MIT Press, 1960.

Romanos, George. *Quine and Analytic Philosophy*. Cambridge, Mass: MIT Press, 1983.

Rorty, Richard. "Response to Davidson" *Rorty and His Critics*. Ed Robert Brandom. Oxford: Blackwell Publishing, 2000. 74-80.

Setiya, Kieran. "Does Moral Theory Corrupt Youth?" *Philosophical Topics*, Forthcoming.

Strawson, P. F. "Two Conceptions of Philosophy." Barrett and Gibson 1990, 310-318.

Stroud, Barry. *The Significance of Philosophical Skepticism*. Oxford: Oxford UP, 1984.

Tennant, Neil. "Radical Interpretation, Logic, and Conceptual Schemes." de Caro 1999, 71-94.

Tersman, Folke. "Davidson and Quine's Empiricism." Kotatko et al 2001, 269-284.

Wittgenstein, Ludwig. *Philosophical Investigations*. (PI) 1951. Trans. G.E.M. Anscombe. Oxford: Basil Blackwell, 2001.

### **Afterword: A Note on Philosophy and History**

According to Alasdair MacIntyre, “Quine has joked that there are two sorts of people interested in philosophy, those interested in philosophy and those interested in the history of philosophy.”<sup>1</sup> Like MacIntyre, I think that this is a mistake. Philosophy cannot be separated from its history. Our philosophical questions are framed in time, using the language of our current moment, and so, are already historical. Attending to historical figures does not merely yield potential answers to the philosophical problems in which one is interested; it shows why those problems *became* problems, and thus, why they have the shapes that they do.

Twentieth-century analytic philosophy is sometimes characterized by its emphasis on Frege’s new logic. Attending to conceptions of logic from recent history, as I have done in this dissertation, is thus to reflect on what philosophical methodology was and how what it was has shaped what it now is. Beyond deepening our understanding of Frege, Quine, and Davidson, I hope to have shown that recognizing the centrality of logic to rational inquiry does not entail a monolithic, restrictive philosophical methodology. By acknowledging the intelligibility of alternate, but nonetheless logical, approaches to philosophical problems, we free ourselves to adopt whichever approach yields satisfying answers.

---

<sup>1</sup> “The Relationship of Philosophy to Its Past,” *Philosophy in History*, Eds. Richard Rorty, J. B. Schneewind and Quentin Skinner, Cambridge: Cambridge UP, 1984. MacIntyre’s “counter-joke” is that, under this conception, those who do philosophy today will be studied only by those doing history of philosophy tomorrow. To the extent that the original joke is meant to malign historians, the joker must abashedly acknowledge her future philosophical irrelevance.